Comparison of system identification methods applied to analysis of inter-area modes

Kaur Tuttelberg, Jako Kilter, and Kjetil Uhlen

***Abstract*—This paper analyses and compares the applicability of various system identification techniques for modal analysis of a multi-area power system. The paper considers system identification applied on PMU measurements of frequency and active power to find a linear multi-input multi-output dynamic model of the primary frequency control of the power system. The multiple input–output pairs correspond to areas of the power system, enabling analysis of inter-area modes. The frequencies and damping ratios of inter-area modes obtained from the identified models are compared to the results of conventional modal analysis (i.e. small signal stability analysis of a linearised power system). Different system identification techniques are compared on simulated wide area PMU measurement data in order to determine the most suitable method for possible on-line analyses.**

***Keywords*: Inter-area oscillations, Modal analysis, System identification, Wide area monitoring**

I. Introduction

Information about critical modes and their damping ratios is a defining question in operating a large power system. In recent years, a number of novel real-time applications have been developed for the task, mostly based on wide area measurement systems; however, the different methods still come with their limitations and are being improved [1]. This paper analyses another approach based on wide area measurements: to apply system identification on a power system. In this case, a linear dynamic model is fitted to the observed changes in quantities and the model is then analysed to determine different properties of the dynamics of the system. In the context of this paper, analysis of inter-area modes based on such an identified model is studied.

System identification has been used before as an alternative way of performing modal analysis on simulated models. It has been suggested for use in cases when only certain modes are of interest, e.g. for PSS tuning [2], [3], but also suggested to overcome some limitations of conventional modal analysis tools [4], [5]. The concept has also been applied when the transfer functions themselves have been of particular interest [6], [7]. With the introduction of wide area measurements enabled by PMUs, the possibility of applying system identification on measured data from the real system has already been studied to some extent as well [8], [9], [10]. Identification of governor models has been demonstrated on real PMU measurement data [11]. Also, methods of model order selection in similar solutions have been analysed [12].

Applying system identification entails a set of problems that have to be considered: mainly providing a suitable set of data, choosing the most appropriate identification method, and specifying the structure of the model. While the fitting algorithm may find a solution for any given dataset, the validity of the model has to be verified as well. In off-line simulations, all of these problems can be solved by hand if necessary, but on-line monitoring would require more automated and robust solutions. This paper compares different system identification techniques to find out which ones would be the most suitable and robust for on-line monitoring applications. The paper concentrates mostly on the choice of identification method and model order selection; less on data selection and model structure in terms of defining inputs and outputs.

This paper applies system identification in terms of a multi-input multi-output black box model, where each input–output pair corresponds to changes in load and frequency of an area of the power system. The paper only considers a very simple power system, where it is trivial to select a set of input and output data for the purpose of system identification. In a more complex system, the selection of inputs and outputs for a meaningful model is a problem by itself and is not treated in this paper. The obtained dynamic model can be seen as a simplified model of the dynamics of primary frequency control and the state of the model as a simplified estimate of the dynamic state of the power system. From the identified model, the frequency and damping ratio of the inter-area mode are extracted.

The implementation in this paper relies on pre-built tools included in the Matlab System Identification toolbox. Thus, the analysis is limited to the transfer function, state space model, and polynomial model estimation functions implemented in this toolbox. This covers a large variety of different possible models, however, many other methods are available for the identification of dynamic systems, e.g. time-domain vector fitting, neural network based learning systems, etc. The paper shows which ones of the classical system identification methods could be applicable in a more automated on-line monitoring application.

The study is based on the well known Kundur two area power system simulated in Digsilent Powerfactory. Based on simulated time series, models are fitted to the input–output data with different system identification methods implemented

This work has been supported by the Estonian TSO, Elering.

K. Tuttelberg and J. Kilter are with the Department of Electrical Power Engineering, Tallinn University of Technology, Tallinn, Estonia (e-mail of corresponding author: kaur.tuttelberg@ttu.ee).

K. Uhlen is with the Department of Electric Power Engineering, Norwegian University of Science and Technology, Trondheim, Norway.

Paper submitted to the International Conference on Power Systems Transients (IPST2017) in Seoul, South Korea June 26–29, 2017

in Matlab. Results of modal analysis reported by Powerfactory are compared to results obtained from the identified models. The results demonstrate that certain system identification methods are more consistent and robust than others, making them more suitable for automated analyses. The comparison of estimated modes indicates that inter-area modes can be detected with the model structure assumed in this paper.

The theoretical basis of the studied method is given in section II, with the specifics of applied system identification methods outlined in section III. Test calculations on the modelled power system are presented in section IV and discussed in section V. Conclusions are given in section VI.

II. Theoretical Background

Primary frequency control is carried out at the turbine–generator unit level. The dynamics of frequency control can be described by a set of differential equations, which in turn can be modelled as a control system. The main components of a unit—the governor, the turbine, and the generator—are modelled by corresponding transfer functions. With a set of simplifications, an area of a power system (or an entire system) can be modelled similarly.

A simplified control system modeling the dynamics of a two area system is depicted in Fig. 1. Here the transfer functions $H_{GTi}(s) = H_{Gi}(s)H_{Ti}(s)$ for any area i model the governor–turbine systems summed as parallel branches and lumped together by evaluating an equivalent droop R_i. $H_{Si}(s)$ are the inertia of all rotating machines (and the frequency dependence of load) lumped into single area (or system) blocks. $H_{L12}(s)$ is the tie-line element, ΔP_{Ri} are the changes in power set-point values, ΔP_{Li} changes in load of each area, and Δf_i changes in frequencies [13].

This model can be simplified further when we analyse only primary frequency control, i.e. model the dynamics before any secondary control is issued. In this case the power set-point values of generators remain unchanged and the corresponding inputs in the control system can be disregarded. In this simplified analysis of the dynamics of secondary frequency control, the multi-area system becomes a multi-input multi-output system with load changes as inputs and frequency changes as outputs.

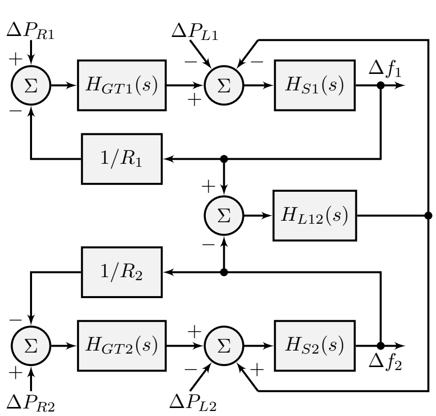

Fig. 1: Dynamic model of the frequency control of a two-area power system.

The described model treats each area as a single node with aggregated load, generation, and control loops and a unified value of frequency. In such a model it is important to define areas that on a system level can be aggregated. The described treatment is also dependent on the possibility of analysing the system in a period of time when the power set-points of all of the generators remain unchanged or change very little (i.e. $\forall i : \Delta P_{Ri} \approx 0$).

III. System Identification

A. Input–Output Data

For a power system that has been divided into a number of appropriately defined areas, a simplified control system modelling the dynamics of its primary frequency control can be identified if sufficient measurement data is available. In other words, a model describing the relationships between changes in load in each of the areas and the corresponding frequency deviations can be fitted to a suitable set of measurement data. This paper analyses the application of various system identification techniques in estimating such fitted models from ambient measurements.

The outputs of the system, i.e. changes in frequency, are simple to measure with PMUs in principle. However, since the areas of the system are lumped into single nodes, some considerations have to be made in order to find the most representative value for the frequency of an area. The larger the areas are, the larger the differences in frequencies of single nodes can be. Information about the structure of the system should be used to combine a set of frequency measurements that best represent the frequency of the area.

The inputs of the system, i.e. the changes in load, are significantly more difficult to measure with PMUs as load is so widely dispersed. On a transmission system level, the measurement of load feeders may become feasible in a not too distant future, but is not something that can be expected at present. This means that more common PMU measurements have to be used to approximate the changes in load. For a practical implementation of the method studied in this paper, a sufficient number of PMUs are required on transmission lines and generators participating in primary frequency control.

This paper suggests two possible approximations. The first would be to estimate the changes in load from the changes in power produced by the generators that perform the significant part of primary frequency control and the power transmitted between areas. This means assuming that when small changes over time are considered, the changes in load and generation are sufficiently close to each other. The approximate change in load in area i would thus be

$$\Delta P_{Li} \cong \sum_j \Delta P_{Gji} + \sum_k \Delta P_{Tki} \tag{1}$$

where ΔP_{Gji} is the change in output power of the jth generator (or a group of generators) participating in primary frequency control in area i and ΔP_{Tki} is the change in power transmitted from the kth to the ith area.

The second option is to approximate the loads as power balances of substations. If a substation in the transmission system has no generation capacity, the load in this node can be estimated as the sum of power on all of the lines coming into the substation, considering power flow directions. If power is also generated at the node, the generation is added to the power balance.

$$\Delta P_{Ln} \cong \sum_{j} \Delta P_{Gjn} + \sum_{k} \Delta P_{Fkn} \tag{2}$$

where ΔP_{Fkn} is the change in power flow from node k to node n and ΔP_{Gjn} is the change in generation in node n.

The proposed method assumes the availability of certain PMU measurements. Regardless of which approximation is used for changes in load, it is necessary to have PMU measurements from the generators that perform the significant part of primary frequency control. The first approximation for load changes requires that the power flows on tie-lines connecting the areas can be measured. For the second approximation of load changes, it is necessary that power flows on all transmission lines in the area can be determined. This does not necessarily require PMUs on all lines, but can also be computed from voltage magnitude and angle differences.

B. System Identification Methods and Model Order

The system identification methods tested in this paper are based on the built-in functions of the System Identification Toolbox in Matlab. The different methods include transfer function estimation by `tfest`, state space model estimation by `ssest` and `n4sid`, and polynomial model estimation by `polyest`, which in turn includes ARX (autoregressive with exogenous inputs), ARMAX (autoregressive–moving-average with exogenous inputs), Output-Error (OE), and Box–Jenkins (BJ) models.

The transfer function estimation method finds common transfer functions that have polynomials of orders np and nz as the denominator and numerator, respectively, while the state space model estimation finds common state space models of order nx. The different polynomial models can all be expressed as special cases of one general model. Given a MIMO polynomial model with nu inputs and ny outputs, the input–output relationships for the lth output of the general model can be expressed as

$$\sum_{j=1}^{ny} A_{lj}(q) y_j(t) = \sum_{i=1}^{nu} \frac{B_{li}(q)}{F_{li}(q)} u_i(t - nk_i) + \frac{C_l(q)}{D_l(q)} e_l(t), \tag{3}$$

where A_{lj}, B_{li}, C_l, D_l, and F_{li} are polynomials of orders na, nb, nc, nd, and nf expressed in q^{-1}. The ARX model is obtained when the C_l, D_l, and F_{li} polynomials are equal to one, OE when $A_{lj} = C_l = D_l = 1$, ARMAX when $D_l = F_{li} = 1$, and BJ when $A_{lj} = 1$ [14].

Due to the nature of the system identification problem that was set up, the models have an equal number of inputs and outputs corresponding to the number of areas the system is divided into. It is difficult to choose one fixed value of model order, despite the work that has been done on model order selection. This paper applies system identification that attempts to fit a number of models in a range of orders $n = n_{\min}, \ldots, n_{\max}$ for each dataset and system identification method. With each iterated order, the various variables defining model order are equal, i.e. $n = np = nx = na = nb = nc = nd = nf$. In transfer function estimation (`tfest`) the number of zeros is smaller than the number of poles by one, i.e. $nz = np - 1$. All input–output relationships are symmetric in the sense that in each identification attempt elements of the order matrices are equal. Input–output delay nk is determined with tools provided in Matlab.

While the `tfest` and `ssest` routines estimate continuous time models, `n4sid` and `polyest` provide discrete time models. In order to compare the results of modal analysis, all discrete time models are converted into continuous time models with the `d2c` function. All estimated models are first checked for stability and only stable ones are analysed further. For each data set a number of models of different orders are obtained and analysed.

IV. Test Calculations

A. Test System and Simulations

The simulations were carried out on the well known Kundur two area power system [13], [15], implemented in Digsilent Powerfactory and depicted in Fig. 2. In these simulations, the system was replicated as closely as possible, except for exciter controls, which were replaced by pre-built models available in the software package, and governors, which were changed in order to have increased variety in the control system. Namely, the exciters were implemented as the SEXS models and the governor models HYGOV (generators G1 and G3) and TGOV1 (generators G2 and G4) were used. All parameters of the controllers were kept at default values, except for droop gains set at $R_1 = 0.05$, $R_2 = 0.03$, $R_3 = 0.04$, and $R_4 = 0.05$. A poorly damped inter-area mode was present in the system, however, it was stable under normal load variations; no PSS was implemented.

The datasets used in system identification were generated in time-domain dynamic simulations (RMS simulation in Powerfactory) of the test system with time varying loads. Each test case was run for a time period of 300 s, step changes in loads were introduced every second and the resulting dynamics of the system were simulated. Various datasets of real load data were adapted to form realistic time series of load changes. Simulated load data is presented in Fig. 3. All changes were made in the two load elements specified in the Kundur two area system (L_7 and L_9). Data was sampled with a time-step of 0.01 s from all simulations. For system identification, load changes approximated by (1) were recorded.

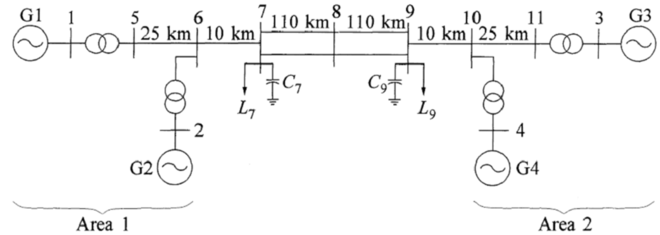

Fig. 2: Schematic of the two area power system used in simulations [13].

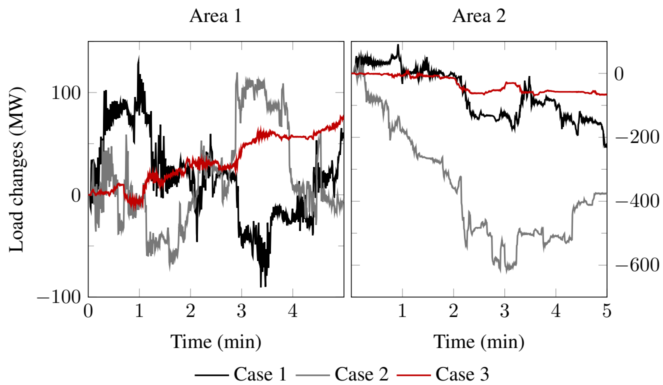

Fig. 3: Cumulative changes in simulated load with respect to initial load. Data from three different simulation cases are plotted.

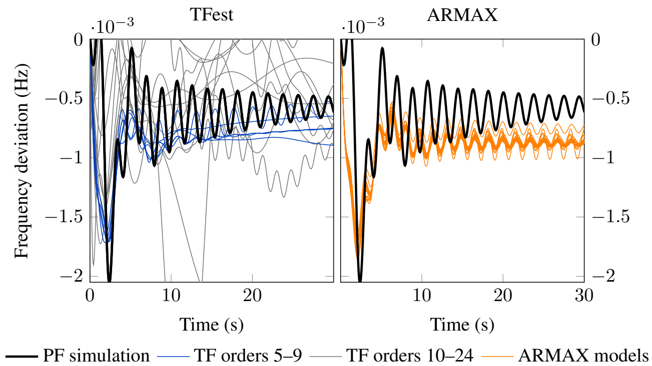

Fig. 4: Step responses of transfer function (TF) models identified with `tfest` and ARMAX polynomial models identified with `armax` from simulation case 3. Step responses of identified models are compared to a step change of load simulated separately in Powerfactory (PF). Frequency changes in area 2 after a step change of load in the same area (normalised to 1 MW) are plotted.

In order to check the accuracy of the identified dynamic models, a separate set of validation data was simulated. For validation, frequency changes after a step change of load (an increase of 50 MW) in each of the areas were simulated and recorded separately. Based on the model used in all of the simulations, small-signal stability analysis was also carried out for the initial state of the system. The inter-area mode was characterised by the poles $-0.086 \pm j4.1$, which corresponds to a frequency of 0.65 Hz and a damping ratio of 0.021.

In the identified MIMO system, two input–output pairs were assumed, corresponding to the two areas. As explained earlier, inputs correspond to load changes and outputs to frequency changes in each of the areas. All identified models were black box without any additional information about the real system provided. Load was estimated as the difference between generation and power exchange between areas according to (1). Area frequencies were taken as the frequencies of buses 5 and 11, which were assumed to be the most central nodes of the two areas. Thus, in total seven measurements were assumed in the studied system: frequencies of buses 5 and 11 and power flows at bus 8 and all four generators.

B. *Identification and Validation of Models*

With the frequency changes and estimated load changes from the three simulation cases, models were identified with all of the methods mentioned in the previous section. All methods were applied to identify models of 20 different orders, ranging from 5 to 24. All iterative methods were set to use 5 iterations. A delay of $nk = 2$ or 0.02 s was identified and used. All identified models were checked for stability, converted to continuous time if necessary, and step responses of stable models compared to validation data. A comparison based on a goodness-of-fit criterion was attempted with NRMSE (normalised root mean squared error), but the criterion failed to identify how well the models mimicked the dynamics of the system. Thus, the models obtained with different identification methods were validated in a visual comparison of step responses in this study.

Fig. 4 presents the comparisons of the simulated step changes and the step responses of the identified models obtained from simulation case 3 with `tfest` and `armax` routines. In case of `tfest`, the step responses of the five lower order models are emphasised as these were the best fits. Even though the identified models captured the overall dynamics of frequency changes quite well, the inter-area oscillations were not very well identifiable. In case of `armax`, very consistent results were seen across all stable models.

The validation of models obtained with `ssest`, `n4sid`, `arx`, `oe`, and `bj` was carried out similarly. The state space models identified with `ssest` and `n4sid` were similar to each other and average in quality when compared to ARMAX models. Polynomial models identified with ARX were similar to ARMAX models, but there were significantly larger differences across models of different orders. Output-Error and Box–Jenkins methods performed very poorly in this application, with a small number of stable models and poor validation results. Some results from state space and ARX methods are also presented in the next subsection.

C. *Comparison of System Identification Methods*

Continuing with the examples of transfer function and ARMAX models, Fig. 5 presents the poles of all models

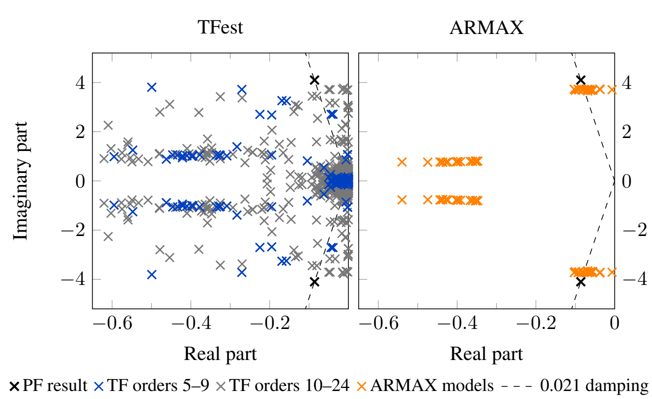

Fig. 5: Poles of transfer function (TF) models and ARMAX polynomial models identified from simulation case 3. Transfer function models of orders 5–9 are plotted in blue, while higher order models are in grey. Poles of estimated models are compared to poles of the inter-area mode identified in Powerfactory (PF). Not all calculated poles fit in the plotted region.

identified with the two methods from simulation case 3. In this study, the poles corresponding to the inter-area mode were identified by assuming an acceptance region for the values. Any pole in the region was assumed to correspond to the inter-area mode. In case of `tfest`, the five lower order models are once again emphasised. Few poles were found close to the expected point determined in modal analysis with Powerfactory. Contrary to that, results obtained with `armax` demonstrate very consistent and good performance. Identified models had poles close to the poles characterising the inter-area mode as found in conventional modal analysis.

The quality of models identified with other methods was similar to the quality seen in validation data. The state space model identification methods `ssest` and `n4sid` performed similarly to each other and provided results of average quality when compared to ARMAX models. Polynomial models identified with ARX had poles similar to the ones found from ARMAX models, but significant differences across model orders were present. Poles of models identified with Output-Error and Box–Jenkins methods were very different from expected correct results.

With all of the identified models, poles close to the correct poles of the inter-area mode were searched for. In most cases, the frequency of the identified mode was close to the expected correct value or at least in the same order of magnitude. In detecting the frequency of the inter-area mode, transfer function estimation with `tfest` performed worst (seen in Fig. 6), while models identified with `armax` gave the best results (seen in Fig. 7). Most of the frequency values obtained with the other methods were within $\pm 10\%$ of the correct value.

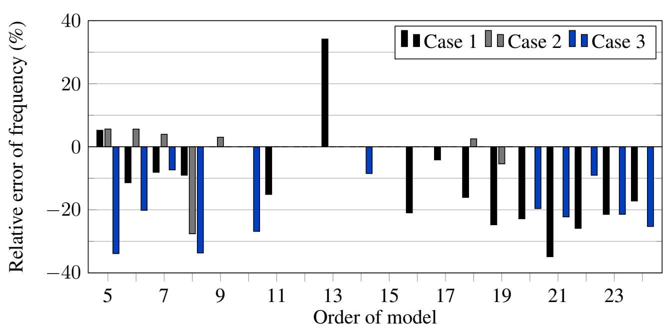

Fig. 6: Relative error of the inter-area mode frequency estimated from transfer function models identified from all three simulation cases.

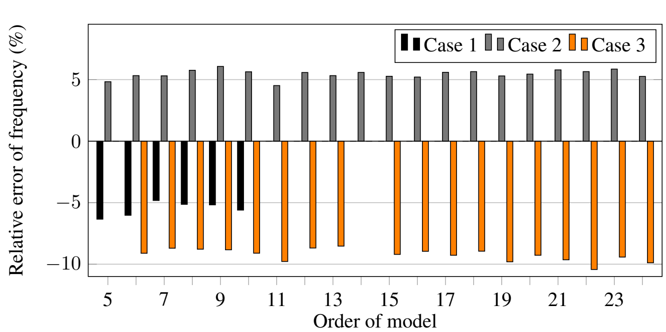

Fig. 7: Relative error of the inter-area mode frequency estimated from ARMAX polynomial models identified from all three simulation cases.

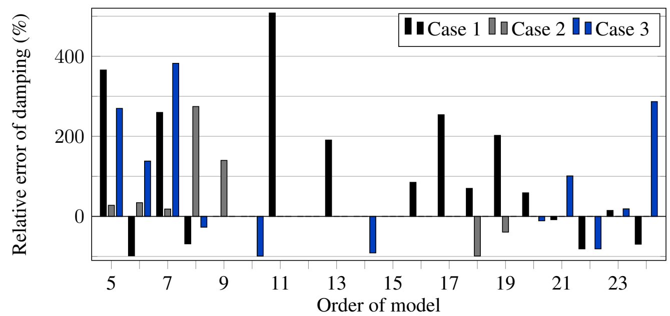

Fig. 8: Relative error of the inter-area mode damping ratio estimated from transfer function models identified from all three simulation cases.

From the identified models, the damping ratio of the inter-area mode was also estimated. In terms of estimating the damping ratio, the identified models performed significantly worse than in estimating the frequency of the mode. Like before, Output-Error and Box–Jenkins models did not provide good results. Both transfer function and state space model estimation provided largely varying results, seen in Fig. 8 and Fig. 9. While a few select models gave a good estimate, most models deviated greatly from expected values. Results of ARX models indicated an increase in accuracy with increasing model orders, seen in Fig. 10.

The accuracy of damping ratios estimated from ARMAX models can be seen in Fig. 11. Many of the ARMAX models provided estimates close to the correct value, while in general they tended to slightly underestimate the damping ratio. Overall the performance of ARMAX model identification was evidently better than that of the other methods. Averaging over

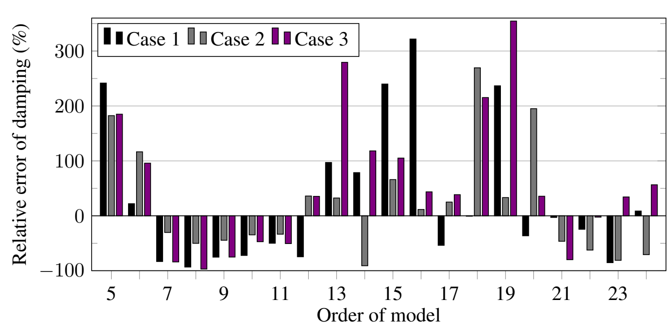

Fig. 9: Relative error of the inter-area mode damping ratio estimated from `n4sid` state space models identified from all three simulation cases.

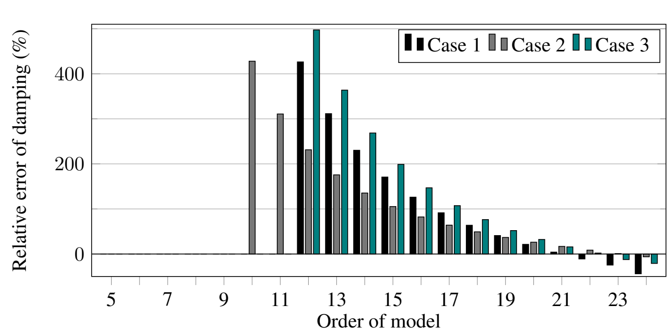

Fig. 10: Relative error of the inter-area mode damping ratio estimated from ARX polynomial models identified from all three simulation cases.

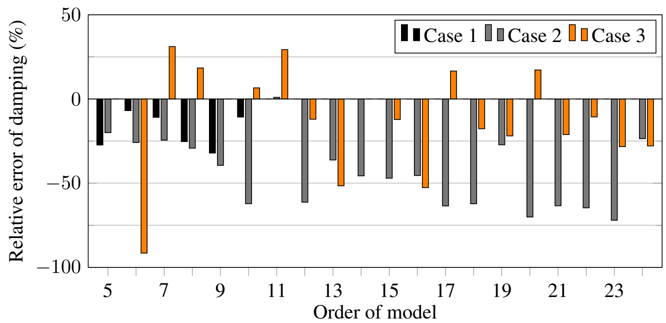

Fig. 11: Relative error of the inter-area mode damping ratio estimated from ARMAX polynomial models identified from all three simulation cases.

the results of all stable ARMAX models of various orders of each case gave damping ratios of 0.017, 0.012, and 0.018 in the three cases (or –19%, –43%, and –14% relative error).

V. Discussion

Out of all the models found with different identification methods, different model orders, and based on different simulation cases, several gave accurate estimates of the frequency and damping ratio of the inter-area mode, while many also failed. However, it was evident that ARMAX polynomial models gave the best and most consistent approximations of the dynamics of the system across various model orders and simulation data. With any of the other studied methods, it would be hard to decide which models were good without validation. With ARMAX models, several similar and good models were obtained, providing redundancy and the ability to check the quality of models by comparing them to each other if good validation data would not be available.

The goal of the presented study was to determine the most suitable system identification method for the described approach to analysing inter-area modes. The paper also demonstrates that the monitoring of inter-area modes based on system identification is feasible, however, many questions remain open. The analysis was based on noiseless data sampled directly from the simulation, not real PMU measurements with limited accuracy. With noisy data, the problems of fitting to noise with higher model orders and additional filtering would also have to be considered. The solution applied in this paper assumed measurement periods where secondary frequency control was not issued or could be considered negligible. To make this assumption valid, additional data about generation set-point changes in the system are needed.

VI. Conclusions

The paper demonstrated the analysis of inter-area modes based on system identification. The aim of the study was to compare the different available system identification methods in order to find the most suitable one for the type of problem analysed in the paper. The study was based on the Kundur two area power system simulated in Digsilent Powerfactory. Three simulation cases were run, where load changed in time, attempting to mimic load variations in a realistic power system. Seven different system identification methods from the System Identification Toolbox in Matlab were applied to fit models to the simulated data. Identified models were analysed to determine the frequency and damping ratio of the inter-area mode, which were also compared to the result found by conventional modal analysis.

The paper identified ARMAX polynomial models as the best suited system identification method for this problem. The paper also demonstrated the feasibility of monitoring inter-area modes based on system identification. With an increasing number of installed PMUs, system identification could be applied to estimate approximate dynamic models and dynamic states of the power system, enabling the analysis of critical modes and other parameters of the system. The approach could be applied in on-line monitoring of modes, PSS tuning, etc. Additional work is required to determine the practical applicability of such monitoring methods. Future work could also analyse alternative dynamic model identification methods and using the identified models to extract other parameters of the dynamics of the system.

References

[1] T. Jiang, L. Bai, G. Li, H. Jia, Q. Hu, and H. Yuan, "Estimating inter-area dominant oscillation mode in bulk power grid using multi-channel continuous wavelet transform," *Journal of Modern Power Systems and Clean Energy*, vol. 4, no. 3, pp. 394–405, 2016.

[2] I. Kamwa and L. Gerin-Lajoie, "State-space system identification-toward MIMO models for modal analysis and optimization of bulk power systems," *IEEE Transactions on Power Systems*, vol. 15, no. 1, pp. 326–335, 2000.

[3] I. Kamwa, "Using MIMO system identification for modal analysis and global stabilization of large power systems," in *Power Engineering Society Summer Meeting, 2000. IEEE*, vol. 2, 2000, pp. 817–822 vol. 2.

[4] J. J. Sanchez-Gasca, K. Clark, N. W. Miller, H. Okamoto, A. Kurita, and J. H. Chow, "Identifying linear models from time domain simulations," *IEEE Computer Applications in Power*, vol. 10, no. 2, pp. 26–30, 1997.

[5] H. Okamoto, A. Kurita, J. J. Sanchez-Gasca, K. Clark, N. W. Miller, and J. H. Chow, "Identification of equivalent linear power system models from electromagnetic transient time domain simulations using Prony's method," in *Decision and Control, 1996., Proceedings of the 35th IEEE Conference on*, vol. 4, 1996, pp. 3857–3863 vol.4.

[6] J. R. Smith, F. Fatehi, C. S. Woods, J. F. Hauer, and D. J. Trudnowski, "Transfer function identification in power system applications," *IEEE Transactions on Power Systems*, vol. 8, no. 3, pp. 1282–1290, Aug 1993.

[7] K. E. Bollinger and W. E. Norum, "Time series identification of interarea and local generator resonant modes," *IEEE Transactions on Power Systems*, vol. 10, no. 1, pp. 273–279, 1995.

[8] J. W. Pierre, D. Trudnowski, M. Donnelly, N. Zhou, F. K. Tuffner, and L. Dosiek, "Overview of system identification for power systems from measured responses," *IFAC Proceedings Volumes*, vol. 45, no. 16, pp. 989 – 1000, 2012.

[9] L. Dosiek and J. W. Pierre, "Estimating electromechanical modes and mode shapes using the multichannel ARMAX model," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1950–1959, 2013.

[10] N. Zhou, Z. Huang, L. Dosiek, D. Trudnowski, and J. W. Pierre, "Electromechanical mode shape estimation based on transfer function identification using PMU measurements," in *2009 IEEE Power Energy Society General Meeting*, 2009, pp. 1–7.

[11] D. T. Duong, K. Uhlen, and E. A. Jansson, "Estimation of hydro turbine-governor system's transfer function from PMU measurements," in *Power and Energy Society General Meeting, 2016 IEEE*, July 2016.

[12] V. S. Peric, T. Bogodorova, A. N. Mete, and L. Vanfretti, "Model order selection for probing-based power system mode estimation," in *Power and Energy Conference at Illinois (PECI), 2015 IEEE*, 2015, pp. 1–5.

[13] P. Kundur, *Power System Stability and Control*. McGraw-Hill, 1993.

[14] *MATLAB and System Identification Toolbox R2016b: User's Guide*, The MathWorks, Inc., Natick, Massachusetts, 2016.

[15] M. Klein, G. J. Rogers, and P. Kundur, "A fundamental study of inter-area oscillations in power systems," *IEEE Transactions on Power Systems*, vol. 6, no. 3, pp. 914–921, Aug 1991.