# A Simulation-Based Classification Approach for Online Prediction of Generator Dynamic Behavior under Multiple Large Disturbances

S. Mahdi Mazhari, *Member, IEEE*, Benyamin Khorramdel, *Student Member, IEEE*, C.Y. Chung, *Fellow, IEEE*, Innocent Kamwa, *Fellow, IEEE,* and Damir Novosel, *Fellow, IEEE*

***Abstract*—This paper proposes a novel method for the machine learning-based online prediction of generator dynamic behavior in large interconnected power systems. Unlike the existing literature in this domain, which assumes faults occur immediately after a steady-state situation, the proposed method takes the possibility of multiple disturbances into account. It is founded on a simulation-based classification approach to indirectly take advantage of phasor measurement unit (PMU) data, which leads to improvements in robustness against load model uncertainties. Relying on offline scenarios, the method developed conducts multiple time-domain simulations (TDSs) in parallel for a set of feasible two-machine dynamic equivalent models (DEMs) for each case. Thereafter, common descriptive statistics are computed for the rotor angles obtained to form the feature space. The values taken via a feature selection process are then applied as inputs to ensemble decision trees, which train models capable of predicting both stability status and generator grouping ahead of time. In online situations, PMU data are used to create DEMs and the predictors are collected by performing parallel TDSs for DEMs. The functionality of the proposed hybrid machine learning and TDS-based approach is verified on several IEEE test systems, followed by a discussion of results.**

***Index Terms*—Feature selection, generator grouping, machine learning, multiple disturbances, phasor measurement unit (PMU), power system dynamic behavior, simulation-based classification (SBC), transient stability.**

## NOTATION

The notation used throughout this paper is reproduced below for quick reference.

*Sets:*

| | |
|---|---|
| $\Omega^{G}$ | Set of generator buses; |
| $\Omega^{SC}$ | Set of offline scenarios; |
| $\Omega_l^{SC}$ | Set of all faults occurred at line $l$; |
| $\Omega^{CG}$ | Set of all unique coherency layouts appeared in $\Omega^{SC}$; |
| $\overline{\Omega_{\boldsymbol{AB}}^{CG}}$ | Set of all unique combinations of the two-cluster generators; |
| $\Omega_{\boldsymbol{AB},l}^{RT}$ | Set of DEMs required for real-time analysis of faults occurred at line $l$; |

*Constants:*

| | |
|---|---|
| $\gamma$ | A cut-off value for rotor angle difference among generators; |
| $\delta_i^j(t)$ | Rotor angle of generator $j$ at instant $t$ for scenario $i$; |
| $n_i^{CG}$ | Number of coherent groups in pattern $i$; |
| $\boldsymbol{Y}_r$ | Admittance matrix of the reduced network; |
| $D_i$ | Damping constant of generator $i$; |
| $H_i$ | Inertia constant of generator $i$; |
| $\delta_i$ | Rotor angle of generator $i$; |
| $x_{di}, x_{qi}$ | $d$- and $q$-axis synchronous reactances of generator $i$; |

*Functions:*

| | |
|---|---|
| $\psi_i(t)$ | Stability index of scenario $i$ at instant $t$; |
| $\overline{\Delta\delta_i}(t)$ | Maximum rotor angle deviation between any pair of generators at instant $t$ for scenario $i$; |
| $\mathcal{D}_{ij}$ | Distance between rotor angles of scenario $i$ for DEM $j$; |
| $\mathcal{R}_{ij}$ | Correlation between rotor angles of scenario $i$ for DEM $j$; |
| $\mathcal{C}_{ij}$ | Crest factor of scenario $i$ for DEM $j$; |
| $\mathcal{MI}(.)$ | Conditional mutual information; |
| $p(.)$ | Probability distribution function (PDF). |

## I. INTRODUCTION

POWER systems are typically confronted by various contingencies that threaten network security. To better address such issues, operators put a set of conservative and preventive considerations into practice to guarantee continuity of supply and avoid instability under high-stress events. However, such strict procedures limit the optimal utilization of network equipment, which consequently increases system operation and planning costs [1]. On the other hand, operators cannot contemplate an infinite set of possible disturbances due to computational and economic restrictions, which means the chance of widespread blackouts remains for some rare scenarios.

Although the issues discussed have garnered the attention of several researchers in the past due to their importance [2], the inclusion of wide-area measurement systems has opened up opportunities to realize novel solution approaches to address operational challenges. For instance, approaches based on online monitoring and event detection [3], fault locating [4], and stability prediction [5] models have been reported in recent years.

In the stability context, the main idea is to develop a platform that can predict network instability with a reasonable time so that operators can trigger corrective and emergency control strategies to maintain synchronism of the system [5]. To this end, several methods have been proposed in recent years, relying on time-domain simulations (TDSs) [6], transient energy functions [7], and machine learning (ML) techniques [8]. Among them, ML-based approaches have gained more attention as a result of their advantages for real-time applications. These methods employ supervised/unsupervised learning frameworks to train stability prediction models using a large set of data collected by offline simulations. Successful implementation of neural networks [8], extreme learning machine [9]–[10], and decision tree (DT) [11]–[12] have been reported in the literature.

The revised manuscript sent May 28, 2020. This work was supported in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada and the Saskatchewan Power Corporation (SaskPower).

S.M. Mazhari, B. Khorramdel, and C.Y. Chung are with the Department of Electrical and Computer Engineering, University of Saskatchewan, Saskatoon, SK, S7N 5A9, Canada (e-mail: s.m.mazhari@usask.ca; bek067@usask.ca; c.y.chung@usask.ca).

I. Kamwa is with the Hydro-Québec/IREQ, Power System and Mathematics Varennes, QC, J3X 1S1, Canada (e-mail: kamwa.innocent@ireq.ca).

Damir Novosel is with the Quanta Technology, Raleigh, NC 27607 USA (e-mail: dnovosel@quanta-technology.com).

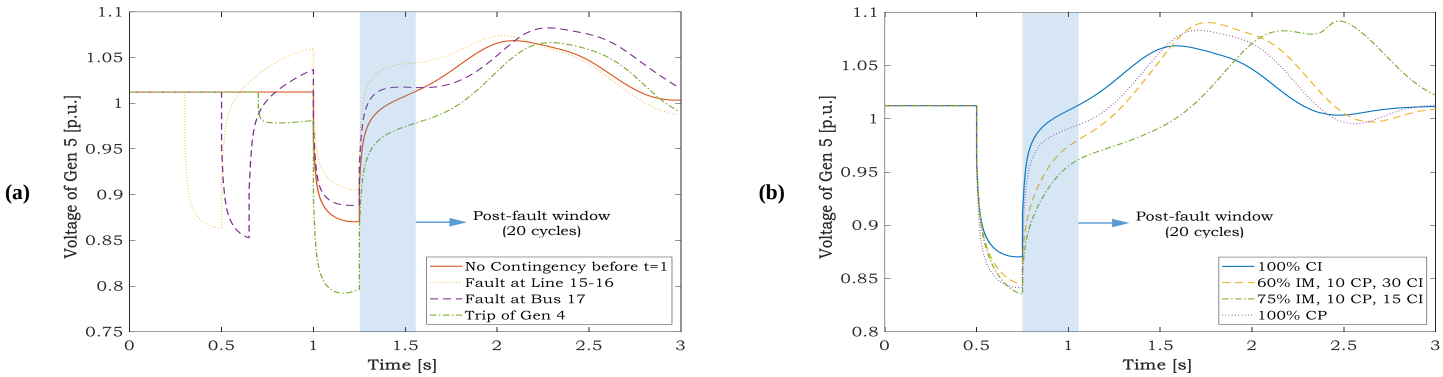


Fig. 1. Voltage values of Gen 5 of the IEEE 39-bus test system under different (a) multiple disturbances, (b) combinations of load models.

While [8]–[12] describe the stability status of the network, they are incapable of presenting any extra information about the system dynamic behavior. In an unstable multi-machine power system, some generators exhibit similar dynamic motion, i.e., their swing curves are so close that they can be considered as single clusters, known as coherent groups, with respect to other machines. Because coherent groups play a significant role in electing a proper post-fault control action [13], prediction of generator dynamic behavior that considers both stability status and generator grouping has been introduced in the specialized literature [14]–[20], and is the main focus of this paper.

In [14], a speed acceleration criterion is developed for generator grouping with a major focus on low coherency events. Although the introduced method is faster than conventional approaches, it still takes several seconds. The problem is formed in a multi-class classification theme in [15]–[19]. In [15], a degree of coherency is defined for each pair of generators and then these are applied as predictors to a DT-based classifier. Post-fault rotor angles obtained from phasor measurement units (PMUs) are employed in [17] and various prediction engines, including DT, random forest, and support vector machine, are used for the model training stage. While [14]–[16] solely focus on identifying generator groups, past studies also address critical generators [17] and the sequence in which generators lose synchronism [18]–[19]. Furthermore, a regression-based framework is performed in [20], where rotor angle swings are forecasted for different post-fault scenarios. Although the predicted values make this method superior to [15]–[19] in terms of better inputs for the decision-making process, topology changes are not considered in [20], which dramatically decreases the prediction accuracy [19].

To date, the majority of existing stability prediction approaches have been developed based on a simplifying assumption in which a fault occurs immediately after a steady-state situation [5]–[20]. However, power systems frequently face various consecutive small and large disturbances that may occur before the system reaches a steady-state condition. According to the North American Electric Reliability Corporation, major disturbances normally include multiple events [21]. A large load change or a component outage that may take place before a fault or contingencies such as successive or simultaneous faults [22], etc., can be considered as multiple large disturbances that challenge past methods in terms of practical applications.

To put it simply, multiple disturbances may occur and, once they do, the post-fault values received from PMUs may not lead to suitable outcomes if injected into the past ML-based methods such as [8]–[20]. For illustration purposes, a three-phase fault is applied to line 21–22 of the IEEE 39-bus test system at $t = 1$. The fault is cleared after 0.25 s and circuit breakers at both ends of the line trip; the voltage values recorded at generator 5 (Gen 5) are shown in Fig. 1(a) (solid line). Three exemplary contingencies are applied before the main fault and the obtained values are depicted in Fig. 1(a). The figure shows the system may not be in a steady state when the last contingency takes place at $t = 1$. The during-fault and post-fault values ($t > 1$) related to multiple disturbances are remarkably different than the case with a single contingency. In past ML-based techniques [8]–[20], all offline simulations, as well as the training stage, were conducted based on such a simplification. Therefore, during-fault and post-fault values, which are used as predictors, may not perform well in all situations, i.e., the data received from PMUs may not exactly follow the solid line shown in the post-fault window of Fig. 1(a). Hence, a framework that can address this deficiency would be of interest to power system operators.

The majority of past stability prediction studies use a constant impedance load model [5]–[20]. While characteristics of large accumulated loads may not change substantially in a short period of time [23], they may experience noticeable seasonal shifts [23]–[24]. Although such variations in load models may not necessarily affect the stability status of the network for a specific contingency [25], they will definitely alter measured post-fault values. Thus, past ML-based methods that directly use PMU data as predictors might be too sensitive to actual load models. As an example, a three-phase fault is applied to line 21–22 of the IEEE 39-bus test system at $t = 0.5$ and cleared after 0.25 s. Four different combinations of load models, including constant impedance (CI), constant power (CP), and induction motor (IM), are considered for buses 15, 16, and 21; all remaining loads are solely modeled by CI. The system is stable in all cases and the measured voltages of Gen 5 are reported in Fig. 1(b). The figure shows notable changes in post-fault values that may affect the prediction accuracy of the past ML-based techniques, such as [7]–[20], in real-life situations. Because dynamic load models are hardly identifiable in real time, it might be helpful to develop stability prediction techniques that have less sensitivity to such uncertainties. A simple method would be expanding offline simulations to cover various load models. However, this may not be feasible in large networks as a huge database would be required to cover an enormous number of possible scenarios. Another approach is to develop predictors that are less sensitive to load model variations, which may deserve further investigation.

Aimed at addressing the drawbacks of the ML-based methods proposed for prediction of generator dynamic behavior to date discussed above, a novel approach is put forward. It is based on a hybrid ML and TDS algorithm that generates the required features for the classification stage. The PMU data are used to form a set of predefined two-machine dynamic equivalent models (DEMs) for the network. The predictors, which reflect suitable resistivity against load model uncertainties, are obtained by applying TDS to the DEMs. Moreover, it introduces a framework that takes potential multiple contingencies of the network into consideration. Such a procedure is introduced as simulation-based classification (SBC) and facilitates deriving generalized prediction models capable of processing multiple contingencies relying on a limited number of offline single-contingency scenarios. To the best of our knowledge, this is the first effort to incorporate either SBC or multiple contingencies into the ML-based transient stability prediction problem. The term "multiple disturbances" is used in this paper to express any sequence of dependent and independent contingencies, regardless of any resulting instability or blackout [22]. The proposed method is successfully applied to several IEEE test systems, and the results obtained are compared with existing techniques.

## II. Generators Dynamic Behavior and General Theme of the Proposed Simulation-Based Classification Approach

As reported in the specialized literature, prediction of generator dynamic behavior can be considered a classification problem that can be solved in two consecutive stages, i.e., stability identification and coherency detection [14]–[19]. If the system is determined to be stable, coherent generators are traced with respect to low-frequency oscillations [14]–[15]. In the case of an unstable network, which is the main focus of this paper, generator grouping is determined to find critical generators and subsequently initiate emergency control actions [16]–[19]. The methods engaged for offline detection of generator dynamic behavior and the general theme of the proposed classification approach are described next.

### A. *Identification of Instability and Generator Groupings*

Several indices are introduced in the literature to identify stability status based on generator rotor angles or speeds. Here, the transient stability status is calculated at any instant of time as follows [11]:

$$\psi_i(t) = \frac{\gamma - \overline{\Delta\delta_i}(t)}{\gamma + \overline{\Delta\delta_i}(t)}, \qquad \forall\, i \in \Omega^{SC} \tag{1}$$

$$\overline{\Delta\delta_i}(t) = \max\left(\left|\delta_i^j(t) - \delta_i^k(t)\right|\right), \qquad \forall\, j,k \in \Omega^G \tag{2}$$

where $\delta_i^j(t)$ is the rotor angle of generator $j$ at instant $t$ for scenario $i$; $\overline{\Delta\delta_i}(t)$ represents the maximum rotor angle deviation between any pair of generators at instant $t$ for scenario $i$; $\gamma$ is a cut-off value for rotor angle difference among generators and is set to $360°$ in this paper; $\Omega^{SC}$ and $\Omega^G$ indicate sets of offline scenarios and generator buses, respectively; and $\psi_i(t)$ denotes the stability index of scenario $i$ at instant $t$, in which positive (negative) values indicate a stable (unstable) network.

Generator groupings are identified based on the rotor angle difference between generators. In this case, hierarchical clustering, which is widely used in previous studies, is put into practice [17]–[19]. For this purpose, an agglomerative (bottom-up) strategy is applied to make a hierarchical cluster tree. First, each generator is set as a single cluster, and then iteratively mixed with nearby clusters based on a linkage criterion. The final clusters are obtained by cutting the tree by setting a threshold value for the linkage criterion, considered to be $360°$ in this work. A detailed process for coherency identification via hierarchical clustering is explained in [17]–[19], but any other method can be used without loss of generality.

### B. *The Proposed Simulation-Based Classification Approach*

In the majority of past methods [8]–[19], the data received from PMU devices are directly selected as predictors. This process increases the sensitivity of prediction models to several factors, including load model, PMU noise, delay, and missing data. To address

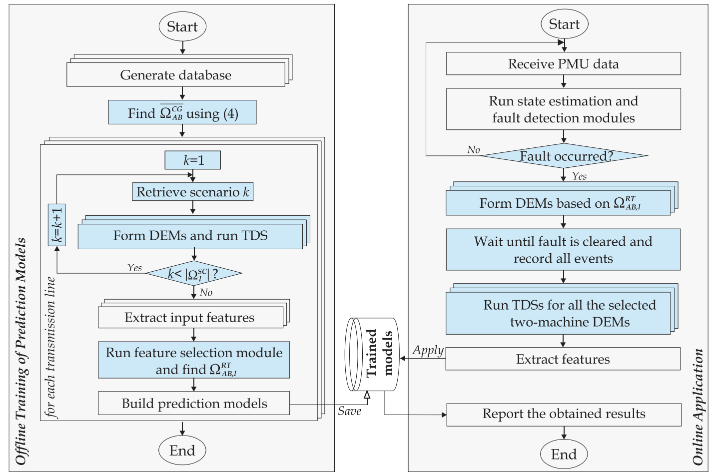


Fig. 2. The overall process of the proposed simulation-based classification approach.

some of these concerns, a simulation-based classification approach is employed in this work so that the predictors are obtained via TDSs. Because running a TDS for a relatively large network in real time is not plausible, a set of possible two-machine DEMs are generated with respect to an offline database and the outputs of online state estimators.

To do so, a set of offline scenarios is generated, $\Omega^{SC}$, and TDS is carried out for each case; all load models are considered as CI in this stage and no multiple contingencies are applied to the database. Then, stability status and generator grouping are determined for each scenario based on Section II.A. Thereafter, a set of all unique coherency layouts appearing in the offline scenarios, $\Omega^{CG}$, is formed. Coherency layouts are then traced and all possible combinations of those patterns that can separate the identified coherent groups into two categories, such as $\boldsymbol{A}$ and $\boldsymbol{B}$, are found ($\Omega_{\boldsymbol{AB},i}^{CG}$). Using statistics, the maximum number of combinations appearing in pattern $i$ would be:

$$\left|\Omega_{\boldsymbol{AB},i}^{CG}\right| = \frac{1}{2}\cdot\left(\binom{n_i^{CG}}{1} + \cdots + \binom{n_i^{CG}}{n_i^{CG}-1}\right) = \frac{2^{n_i^{CG}}-1}{2}, \qquad \forall\, i \in \Omega^{CG}, n_i^{CG} > 1 \tag{3}$$

where $n_i^{CG}$ indicates the number of coherent groups in pattern $i$. Finally, all unique combinations of the two-cluster generators ($\overline{\Omega_{\boldsymbol{AB}}^{CG}}$) are obtained by (4):

$$\overline{\Omega_{\boldsymbol{AB}}^{CG}} = \bigcup_{i \in \Omega^{CG}} \Omega_{\boldsymbol{AB},i}^{CG} \tag{4}$$

For each offline scenario, the equivalent model of generators is calculated for all combinations found in (4) based on the available system data and the steady-state values, as explained in Section III. Parallel TDSs are then carried out to record the dynamic behaviors of the reduced systems. The rotor angles obtained from such simulations are used to extract predictors; because these features are largely obtained via TDSs, rather than instantaneous PMU data, they express less sensitivity against different sources of uncertainties, as shown in Section V. Furthermore, $n_i^{CG}$ is a relatively small number, which makes $\overline{\Omega_{\boldsymbol{AB}}^{CG}}$ a finite set. Simulations empirically show that, for the IEEE 145-bus test system, $n_i^{CG} < 7$, the average $n_i^{CG}$ is 3.24, and the size of $\overline{\Omega_{\boldsymbol{AB}}^{CG}}$ is 16 for the database generated in [19].

To find the prediction models, each transmission line is examined separately. Depending on the system under study, all faults occurring in a sample line $l$, $\Omega_l^{SC}$, only bring about a limited number of coherent groups. Because the faulted line can be easily identified

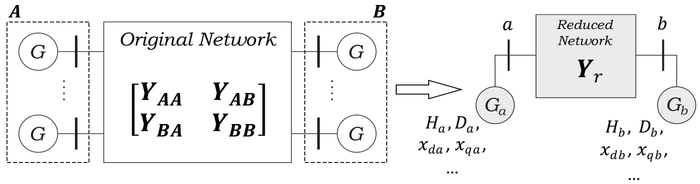


Fig. 3. Network aggregation using a structure preservation technique.

using available event detection algorithms [3]–[4], developing a coherency prediction model for each line, which is rarely addressed in theliterature, can substantially increase the prediction accuracy. In this respect, the feature selection process explained in Section IV.B is employed to decrease the number of DEMs required for real-time analysis ($\Omega_{\boldsymbol{AB},l}^{RT}$) of a particular line, such as $l$.

The overall procedure of the proposed classification approach is illustrated in Fig. 2, where the main elements added to the scheme are highlighted. This diagram shows that, in online applications, the data received from PMUs are injected into the state estimation and event detection modules of the energy management system (EMS), and the important events are recorded. Once a fault occurs in the network and is identified by the detection tools, the first phase of the proposed method is triggered and several DEMs are calculated for the network based on the latest recorded steady-state situation. Once the detection algorithms, digital fault recorders, or relay signals indicate that the fault has been cleared, real-time TDSs are conducted in parallel for multiple two-machine networks in which all events and multiple contingencies are considered during simulations. Multiple contingencies, if there are any, will be applied while running TDSs for each DEM. In other words, if a single contingency occurred, the TDSs only simulate one disturbance; however, if another event has taken place before the system reaches a steady-state situation, the TDSs starts from the last steady-state snapshot and simulates all events occurred in between. The data obtained in the simulations are then applied to the trained classification models, which reveal the dynamic behavior of the system. Because the majority of the input features are calculated based on simulations (and not direct PMU data), the predictors are less sensitive to load model uncertainties and capable of handling multiple contingencies. A detailed procedure for creating DEMs and the proposed solution framework are described in the following sections.

It might be helpful to mention that although the proposed method does not directly apply the PMU data into the prediction phase, it employs PMU data to run state estimation and fault detection modules, as stated in Fig. 2. Hence, PMU placement should be carried out in a way that full observability of the network is met in normal operation. However, no other restrictions, such as placing PMUs on generator buses [9]–[20], are needed in the proposed framework.

## III. Network Aggregation and Two-Machine DEMs

A large multi-area interconnected power system is represented by multiple two-machine DEMs in this work. The last snapshot of the steady-state situation is employed for the aggregation calculations. Based on Section II.B, each element of $\overline{\Omega_{\boldsymbol{AB}}^{CG}}$ contains two sets of generators, such as $\boldsymbol{A}$ and $\boldsymbol{B}$; an equivalent generator of each set is used for TDSs. The processes for finding both nodal and machine equivalents are described next.

### A. Nodal Aggregation

A structure preservation technique, shown in Fig. 3, is conducted for bus reduction [26]–[27] where sets of nodes $\boldsymbol{A}$ and $\boldsymbol{B}$ are respectively portrayed by terminal buses $a$ and $b$ while the following conditions are met:

- Aggregating each set of nodes does not change currents and voltages of the retained nodes;
- The power injection from each terminal node remains the same as from the related aggregated nodes.

Applying this method to both sides, the coherent buses are converted into equivalent nodes using ideal phase-shifting transformers with complex ratios. Assuming nodal equations of the original network, shown in Fig. 3, are as follows:

$$\begin{bmatrix} \boldsymbol{I_A} \\ \boldsymbol{I_B} \end{bmatrix} = \begin{bmatrix} \boldsymbol{Y_{AA}} & \boldsymbol{Y_{AB}} \\ \boldsymbol{Y_{BA}} & \boldsymbol{Y_{BB}} \end{bmatrix} \begin{bmatrix} \boldsymbol{V_A} \\ \boldsymbol{V_B} \end{bmatrix} \tag{5}$$

Then the reduced admittance matrix, $\boldsymbol{Y}_r$, can be calculated by [26]:

$$\boldsymbol{Y}_r = \begin{bmatrix} \boldsymbol{k}_a^{*t} \cdot \boldsymbol{Y_{AA}} \cdot \boldsymbol{k}_a & \boldsymbol{k}_a^{*t} \cdot \boldsymbol{Y_{AB}} \cdot \boldsymbol{k}_b \\ \boldsymbol{k}_b^{*t} \cdot \boldsymbol{Y_{BA}} \cdot \boldsymbol{k}_a & \boldsymbol{k}_b^{*t} \cdot \boldsymbol{Y_{BB}} \cdot \boldsymbol{k}_b \end{bmatrix} \tag{6}$$

$$\boldsymbol{k}_a = \boldsymbol{V_A} \cdot V_a^{-1}, \qquad \boldsymbol{k}_b = \boldsymbol{V_B} \cdot V_b^{-1} \tag{7}$$

$$V_a = \frac{\boldsymbol{V_A} \cdot \boldsymbol{I_A^*}}{\left(\sum_{i=1}^{|\boldsymbol{A}|} \boldsymbol{I_A}(i)\right)^*}, \qquad V_b = \frac{\boldsymbol{V_B} \cdot \boldsymbol{I_B^*}}{\left(\sum_{i=1}^{|\boldsymbol{B}|} \boldsymbol{I_B}(i)\right)^*} \tag{8}$$

where $\boldsymbol{k}_a$ and $\boldsymbol{k}_b$ represent complex ratio vectors of the ideal transformers. Eq. (8) verifies the power balance and unveils voltage values at the terminal buses.

### B. Generator Aggregation

The dynamic equivalent of a group of generating units is known as a single generator reflecting the same speed and power flow parameters as the original generators during any perturbation. Such equivalent generators are connected to terminal buses, shown in Fig. 3, and represent aggregated parameters of the generating units and their control systems. A non-iterative procedure proposed in [26] is employed in this paper to calculate equivalent parameters; it relies on structure preservation of the coefficient matrices in the time-domain representation of synchronous machines. Based on this method, the inertia constant ($H$), damping constant ($D$) on the system MVA base, and rotor angle ($\delta$) of the equivalent machine connected to bus $a$, $G_a$, can be calculated by (9):

$$H_a = \sum_{i=1}^{|\boldsymbol{A}|} H_i\,, \qquad D_a = \sum_{i=1}^{|\boldsymbol{A}|} D_i\,, \qquad \delta_a(0) = \frac{\sum_{i=1}^{|\boldsymbol{A}|} H_i \cdot \delta_i}{H_a} \tag{9}$$

The equivalent $d$- and $q$-axis synchronous reactances on the system MVA base are obtained by (10)–(11). $x_{di}$ and $x_{qi}$ respectively represent the $d$- and $q$-axis synchronous reactances of generator $i$. The same procedure can be used to obtain equivalent transient and subtransient reactances [26].

$$x_{da} = \frac{1}{\sum_{i=1}^{|\boldsymbol{A}|} \boldsymbol{k}_a(i) \left(\frac{\cos^2(\delta_i - \delta_a)}{x_{di}} + \frac{\sin^2(\delta_i - \delta_a)}{x_{qi}}\right)} \tag{10}$$

$$x_{qa} = \frac{1}{\sum_{i=1}^{|\boldsymbol{A}|} \boldsymbol{k}_a(i) \left(\frac{\cos^2(\delta_i - \delta_a)}{x_{qi}} - \frac{\sin^2(\delta_i - \delta_a)}{x_{di}}\right)} \tag{11}$$

Other equivalent parameters including transient time constants, excitation system, and turbine governor are also considered in this work, but not reported here as they are not a part of the contribution of this paper. A detailed process for calculating those values can be found in [26]–[27]; it goes without saying that any other technique can be used without loss of generality.

Dynamic aggregation methods require information about coherent groups that lie behind the proximity of generator speed in each group. Thus, depending on deviation in their speed, some errors can occur in calculating parameters of the equivalent machine. Because $\overline{\Omega_{\boldsymbol{AB}}^{CG}}$ contains all possible two-machine coherent groups, such errors may appear in some DEM layouts depending on the system operating point. However, they rarely affect the performance of the proposed prediction approach because the time-domain solutions would feed into a machine learning procedure before any decision is made.

## IV. The Proposed Solution Framework

Based on $\overline{\Omega_{\boldsymbol{AB}}^{CG}}$ calculated in (4), several two-machine equivalent networks are derived for each offline scenario using the last available steady-state snapshot. Then, a TDS is carried out for each reduced network in which all consecutive faults and contingencies are implemented by updating $\boldsymbol{Y}_r$ or parameters of the equivalent generators during simulations. Overall, $\overline{\Omega_{\boldsymbol{AB}}^{CG}}$ parallel simulations are conducted for each offline scenario and rotor angle values of the equivalent machines are recorded. Those values are then pushed into the feature extraction, feature selection, and training procedures, as explained below.

### A. Feature Extraction

Assume $\delta_{a,ij}$ and $\delta_{b,ij}$ are rotor angle values of the equivalent generators $a$ and $b$, which are calculated for coherency pattern $j$ ($j \in \overline{\Omega_{\boldsymbol{AB}}^{CG}}$) of scenario $i$ ($i \in \Omega^{SC}$). The following common descriptive statistics are employed to extract features from the rotor angles obtained.

#### 1) Distance between Rotor Angles

The sum of the absolute difference between rotor angle samples, $\mathcal{D}_{ij}$, is calculated based on (12) and considered as a feature.

$$\mathcal{D}_{ij} = \sum_{n\in\Omega^n} \left|\delta_{a,ij}(n) - \delta_{b,ij}(n)\right|, \qquad \forall\, i \in \Omega^{SC}, j \in \overline{\Omega_{\boldsymbol{AB}}^{CG}} \tag{12}$$

where $\Omega^n$ is the set of time samples. If coherent generators are correctly selected, a very small amount of deviation between these two signals indicates a stable system, while very large values correspond to an unstable one.

#### 2) Correlation between Rotor Angles

In a stable system, rotor angles of the equivalent generators are supposed to swing with each other. As such, the correlation between these two sets of data, $\mathcal{R}_{ij}$, can be used as a feature to identify the dynamic behavior of the system.

$$\mathcal{R}_{ij} = \frac{\sum_{n\in\Omega^n}\left[\left(\delta_{a,ij}(n) - \overline{\delta_{a,ij}}\right)\cdot\left(\delta_{b,ij}(n) - \overline{\delta_{b,ij}}\right)\right]}{\sqrt{\sum_{n\in\Omega^n}\left(\delta_{a,ij}(n) - \overline{\delta_{a,ij}}\right)^2}\cdot\sqrt{\sum_{n\in\Omega^n}\left(\delta_{b,ij}(n) - \overline{\delta_{b,ij}}\right)^2}} \tag{13}$$

where overbar represents the average value of the signal.

#### 3) Crest Factor

Crest factor is an index of a signal indicating the ratio of peak values to the effective value [28]. This factor represents the extremeness of the peaks in a waveform. If the generators in a two-machine equivalent system are not properly grouped, the rotor angles may have a wavy character, i.e., the stability status of the system may not be clearly identified from the rotor angle signals obtained. Thus, the crest factor is calculated by (14) and considered a feature to reflect this phenomenon.

$$\mathcal{C}_{ij} = \frac{\left\|\delta_{a,ij} - \delta_{b,ij}\right\|_\infty}{\sqrt{\frac{1}{|\Omega^n|}\sum_{n\in\Omega^n}\left(\delta_{a,ij} - \delta_{b,ij}\right)^2}}, \qquad \forall\, i \in \Omega^{SC}, j \in \overline{\Omega_{\boldsymbol{AB}}^{CG}} \tag{14}$$

where $\|\,.\,\|_\infty$ represents the $\ell$-infinity norm.

In addition to the statistical factors explained above, fault type, fault duration, and stability status in the steady-state, $\psi_i(0)$, are used as input features. The first two parameters can be easily obtained from event detection modules while $\psi_i(0)$ is found by (1). These parameters have been shown to increase overall stability prediction accuracy [11].

### B. Feature Selection

Prediction of generator dynamic behavior can be interpreted as solving classification problems for a set of input features and an observation vector. Assume $C = [c_k]_{|\Omega^{SC}|\times 1}$ is the observation vector; in the case of stability prediction, it contains the stability status of each offline scenario ($c_k \in \{0,1\}$) [9]–[12]. Considering $\Omega^{CG}$ as a set of all unique coherency layouts in $\Omega^{SC}$, $C$ reflects the class label of each scenario ($c_k \in \{1, \ldots, |\Omega^{CG}|\}$) for generator grouping prediction [15]–[19]. In this paper, a unique model is trained for each line; so, $C$ is confined to a set of all fault scenarios that occurred in each line.

Considering the features extracted in Section IV.A, the total size of the feature space for each line would be $3 \times \left|\overline{\Omega_{\boldsymbol{AB}}^{CG}}\right|$. Because calculation of the statistical features discussed above does not incur a heavy computational burden, the size of $\overline{\Omega_{\boldsymbol{AB}}^{CG}}$ is the major factor that may confine application of the proposed framework in real-time situations. To resolve this issue, a conditional mutual information ($\mathcal{MI}$)-based method is employed to minimize the two-machine coherency set [29]. $\mathcal{MI}$ is a widely used information-theoretic quantity that measures how much information is communicated, on an average basis, in one random variable about another. For the observation vector $C$ and a feature vector $X$, $\mathcal{MI}$ is defined as [29]:

$$\mathcal{MI}(C;X) = \sum_{c\in C}\sum_{x\in X} p(c,x)\cdot\ \log\left(\frac{p(c,x)}{p(c)\,.\,p(x)}\right) \tag{15}$$

where $p(c,x)$ indicates the joint probability distribution function (PDF) of $C$ and $X$, and $p(c)$ and $p(x)$ respectively represent the PDFs of $C$ and $X$. $c$ and $x$ denote any points belonging to $C$ and $X$, respectively. Notably, a binned format of $p(c,x)$, $p(c)$, and $p(x)$ can be utilized to approximate these variables with a low computational burden [29].

For each two-machine DEM, $\mathcal{MI}$ between the statistical features reported in Section IV.A and the observation vector $C$ is calculated. Among them, the feature with the maximum $\mathcal{MI}$, $X_i$, is selected as representative of that layout:

$$X_i = \arg\max_{F_i\in\{\mathcal{D}_i,\mathcal{R}_i,\mathcal{C}_i\}}\{\mathcal{MI}(F_i,C)\}, \qquad \forall i \in \overline{\Omega_{\boldsymbol{AB}}^{CG}} \tag{16}$$

Assume $X$ is the set of representative features, defined by (16), and $S$ is the subset of selected features. For a new feature $X_i \in X$ to be selected, it is expected that the amount of information about $C$ provided by $X_i$, which is not already supplied by $S$, must be the largest of all candidate features in $X\backslash S$ (all members of $X$ that are not members of $S$). In other words, the conditional mutual information of $C$ and $X_i$, given the subset of already selected features $S$, should be maximized. Such conditional $\mathcal{MI}$ can be calculated as follows:

$$\mathcal{MI}(C;X_i|S) = \mathcal{MI}(C;X_i) - \beta\cdot\sum_{X_s\in S}\mathcal{MI}(X_s;X_i) \tag{17}$$

where $\beta$ is a factor that controls the redundancy penalization among single features and is set to $1/|S|$ in this paper [29].

The overall process of finding the optimal number of two-machine DEMs required for real-time prediction in line $l$, ($\Omega_{\boldsymbol{AB},l}^{RT}$), is as follows:

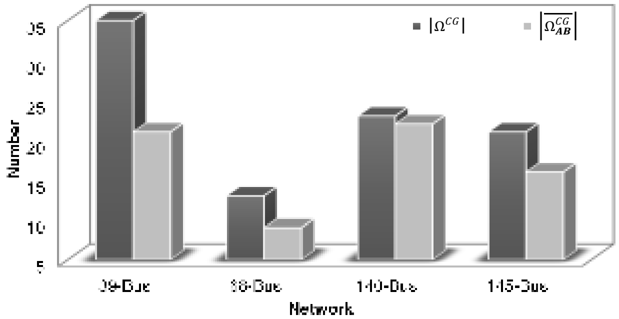


Fig. 4. Coherent patterns* identified in offline scenarios
* Patterns with rare possibility of occurrence ($\leq 3 \times 10^{-4}$) are ignored.

TABLE I
DATA FOR THE NETWORKS USED IN SIMULATIONS

| Network | # of transmission lines | $\lvert\Omega^{G}\rvert$ | $\lvert\Omega^{SC}\rvert$ (unstable %) |
|---|---|---|---|
| 39-Bus | 34 | 10 | 12500 (23.10%) |
| 68-Bus | 66 | 16 | 12500 (13.92%) |
| 140-Bus | 206 | 48 | 20000 (17.06%) |
| 145-Bus | 401 | 50 | 20000 (14.14%) |

Step 1) Find the representative feature of each reduced network using (16);

Step 2) For all features $X_i \in X$, calculate $\mathcal{MI}(C; X_i)$;

Step 3) Find the feature that maximizes $\mathcal{MI}(C; X_i)$ and add it to the subgroup of selected features $S$;

Step 4) Select the next suitable feature, $X^{+}$, by:

$$X^{+} = \arg \max_{X_i \in X \setminus S} \{\mathcal{MI}(C; X_i | S)\} \tag{18}$$

Step 5) Update $S$ and repeat Step 4 until the stopping criterion is met;

Step 6) Consider the obtained $S$ as the set of selected two-machine DEMs for line $i$, $\Omega_{\boldsymbol{AB},i}^{RT}$.

### C. The Training Process

Simulations empirically show that the fault location significantly affects either the system instability or the generator coherency layout. Thus, it is reasonable to train a single prediction model for all faults taking place in each line; this action reduces the search space and increases the prediction accuracy. Because a minimum level of observability is mandatory for the successful operation of the state estimators, there must be enough measurement devices in the network to also handle the fault detection process. Such an assumption will not restrict the application of the proposed method in real-life systems, while a single prediction model can also be derived for the whole network in a straightforward manner.

In this paper, the training stage is handled by ensemble DTs [11]. DT is one of the most popular machine learning tools, and is widely used in the literature for classification and regression purposes. DT has been used in several recent papers in this domain, e.g., [17]–[19], and shown to perform very well with a medium-sized database [11], which is the case in this paper. However, because the ML method, itself is not part of the contribution of this work, any other technique can also be used without loss of generality. Once the set of selected two-machine DEMs is obtained by the proposed conditional $\mathcal{MI}$-based method, the associated features calculated in Section IV.A are used as the input features of the ensemble DTs; in total, $3 \times \left|\Omega_{\boldsymbol{AB},i}^{RT}\right|$ features are fed into each DT. The DTs are constructed based on the standard classification and regression tree (CART) and the ensembles are fabricated using the boosting technique. The trained models are saved on a local disk and retrieved during online applications.

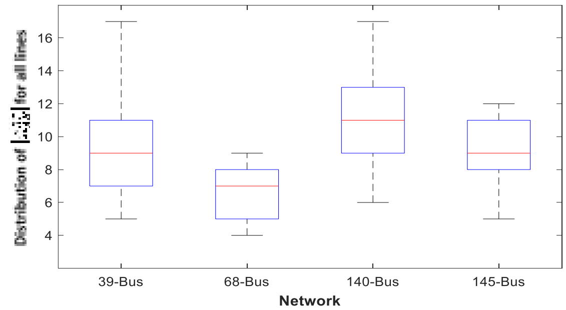


Fig. 5. Distribution of the selected features for different test systems.

## V. TESTS AND RESULTS

The proposed solution framework is implemented in the Python environment, through which an automated offline scenario generation process is formed. The Python scripts interact with the PSS/E software [30] to run transient stability for each scenario, build DEMs, and conduct TDSs for two-machine DEMs. The developed package is tested on several networks including IEEE 39-, 68-, 140-, and 145-bus test systems. The number of offline scenarios generated for each network is shown in Table I; for each network, 3-phase faults are applied to buses and transmission lines of the original system. The location of each fault, number of faults in each line, and fault durations (2–15 cycles) are randomly selected. The fault impedance is set based on [11] and the load consumption at each bus is randomly scaled by 0.65–1.25 with respect to the base demand. Moreover, $N-1$ and $N-2$ situations are considered while generating scenarios and cover about 20% of the dataset to reflect topological changes. The TDS is conducted for 20 s after fault clearance in each case, and the data obtained are recorded in phasor format, two samples per cycle, with respect to the standard PMUs [17]. The computer used in this study featured an Intel 3.4-GHz CPU with 16 GB of RAM. Using the established platform, various simulations were carried out in two scenarios to evaluate the performance of the proposed method, as described next.

### A. First Scenario

In this scenario, the prediction of generator dynamic behavior is conducted for IEEE test systems, and the results obtained are compared with those from existing techniques. The generator grouping algorithm, explained in Section II, is applied to an offline database and the total number of unique coherent patterns identified in each network is shown in Fig. 4. This figure shows the IEEE 39-bus system reflects the greatest number of patterns among the studied networks. Then, each coherent pattern is solely traced, all possible combinations that can divide it into two groups of generators are found, and these are applied to (4). The $\left|\Omega_{\boldsymbol{AB}}^{CG}\right|$, shown in Fig. 4, represents the total number of unmatched combinations of two-cluster generators appearing in the database. It indicates the number of parallel two-machine TDSs that should be carried out during the training process for each offline scenario. For instance, in the case of the IEEE 140-bus system, 22 parallel two-machine DEMs are formed for each scenario and a single TDS is conducted for each.

To reduce the number of two-machine TDSs for online applications, the data obtained are applied to a feature selection process explained in Section IV.B. Based on that, the optimal number of two-machine TDSs is identified for each faulted line, and the results derived are summarized in Fig. 5. Comparing Figs. 4 and 5 shows the feature selection process can substantially reduce the input features in all networks; as an example, the average number of two-machine TDSs is reduced from 21 to 9 in the IEEE 39-bus system, which

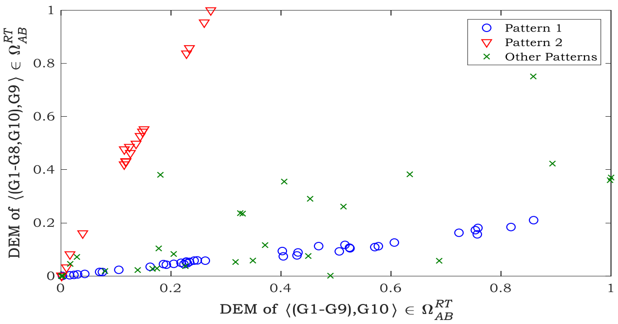


Fig. 6. Distribution of normalized crest factors for two DEMs representing training cases wherein a fault occurred at line 2–3 of the IEEE 39-bus system.

TABLE II
AVERAGE STABILITY ASSESSMENT/PREDICTION ACCURACY FOR THE DATASET (%)

| Network | Accuracy (%) | | | | | |
|---|---|---|---|---|---|---|
| | Assessment based on a properly selected DEM* | | | Prediction based on the proposed method | | |
| | Stable | Unstable | All | Stable | Unstable | All |
| 39-Bus | 94.48 | 85.32 | 89.90 | 99.45 | 97.26 | 98.35 |
| 68-Bus | 95.36 | 86.43 | 90.89 | 99.91 | 96.17 | 98.04 |
| 140-Bus | 94.90 | 85.85 | 90.37 | 99.83 | 98.05 | 98.94 |
| 145-Bus | 93.27 | 84.66 | 88.96 | 99.39 | 96.84 | 98.11 |

* Indicates a single DEM that represents the correct grouping of generators.

TABLE III
COMPARISON OF AVERAGE PREDICTION ACCURACY FOR DIFFERENT IEEE TEST SYSTEMS (%)

| Item | Test situation | 39-bus | 68-bus | 140-bus | 145-bus |
|---|---|---|---|---|---|
| Stability status prediction | [17] with PFC = 20 | 95.59 | 94.39 | 95.58 | 92.67 |
| | [17] with PFC = 60 | 96.96 | 96.36 | 97.50 | 95.47 |
| | [19] with PFC = 60 | 97.45 | 95.72 | 97.75 | 97.32 |
| | Proposed method | **98.35** | **98.04** | **98.94** | **98.11** |
| Coherency prediction | [17] with PFC = 20 | 87.27 | 87.95 | 92.96 | 86.13 |
| | [17] with PFC = 60 | 92.67 | 92.37 | 92.79 | 92.32 |
| | [19] with PFC = 60 | 94.06 | 92.79 | 92.80 | 92.89 |
| | Proposed method | **96.45** | **96.53** | **94.85** | **95.17** |

represents more than a 57% reduction in computational burden in online applications. It might be helpful to mention that although faults applied to some specific lines trigger a single coherent pattern for the original network because $N-1$ and $N-2$ situations are considered in this work, more than one DEM is required for each line to perform the prediction with suitable accuracy.

The network aggregation and the two-machine DEM calculations reported in Section III are the backbones of the feature generation process for the proposed technique. Considering the simplifications applied, such a procedure itself, can assess approximate stability for a reduced network. However, in real-time applications, the generator grouping is not known, and thus the reduced network cannot be formed. Nonetheless, because the proposed framework runs TDSs for all possible DEMs, it might be helpful to evaluate the stability assessment accuracy of the two-machine system, with known coherent groups; it gives a sense of the quality of the input features. The results of such an assessment are shown in Table II. This table shows that, on average, the two-machine DEM could lead to 90.89% accuracy for the IEEE 68-bus system; the lowest prediction accuracy is 88.96% for the IEEE 145-bus test system. The simulation results reveal that the reduced system can fairly predict the stable cases, with above 93% accuracy, but performs moderately on unstable cases. It is empirically seen in simulations that the two-machine system may fail to correctly assess stability for cases with more than two generator groups. Nevertheless, the obtained results indicate that the approximate calculations of the two-machine system have the potential to be used as input features of a stability prediction model.

TABLE IV
AVERAGE PREDICTION ACCURACY FOR DIFFERENT TOPOLOGIES OF THE DATASET (%)

| Network | Stability status prediction | | | Coherency prediction | | |
|---|---|---|---|---|---|---|
| | $N-0$ | $N-1$ | $N-2$ | $N-0$ | $N-1$ | $N-2$ |
| 39-Bus | 98.29 | 98.65 | 99.01 | 96.62 | 95.82 | 95.58 |
| 68-Bus | 98.14 | 97.65 | 97.24 | 96.71 | 95.85 | 95.76 |
| 140-Bus | 99.07 | 98.56 | 97.35 | 95.02 | 94.50 | 93.18 |
| 145-Bus | 98.16 | 98.02 | 97.25 | 95.30 | 94.71 | 94.40 |

TABLE V
SENSITIVITY ANALYSIS OF GENERATOR MODELING ERROR ON AVERAGE PREDICTION ACCURACY FOR THE IEEE 39-BUS SYSTEM

| # | Subtransient reactance | Subtransient time constant | Excitation system | Stability status prediction (%) | Coherency prediction (%) |
|---|---|---|---|---|---|
| 1 | -* | - | - | 88.19 | 80.71 |
| 2 | ✓+ | - | - | 91.41 | 84.54 |
| 3 | ✓ | ✓ | - | 92.20 | 85.16 |
| 4 | ✓ | - | ✓ | 97.32 | 95.53 |
| 5 | ✓ | ✓ | ✓ | 98.35 | 96.45 |

* Average value is used for equivalent modeling; + Detailed value is used.

The proposed solution framework, shown in Fig. 2, is applied to each network to solve both stability prediction and generator grouping problems. A stratified 5-fold technique is employed to divide the whole database; the evaluation process is reiterated five times utilizing diverse training sets, and 20% of the dataset is used as test samples in each iteration. The average stability prediction results obtained are reported in Table II, which shows the lowest prediction accuracy obtained on the dataset is 98.04%. Moreover, the proposed method is capable of correctly predicting more than 99.39% of stable cases, i.e., it is unlikely to trigger an incorrect emergency control action for a stable system.

For the sake of comparison, the techniques reported in [17] and [19] are implemented and applied to the same database; as these methods require post-fault data, different post-fault cycles (PFC) of voltage and rotor samples are considered input features for the classification process. It might be helpful to mention that PFC is a period for which the prediction algorithm waits to receive enough amount of data before making the prediction; for a 60-Hz system, PFC=60 means the prediction model waits for 1 second after clearing the fault. The results obtained are reported in Table III, which shows the proposed method outperforms the existing techniques in all networks. For instance, in case of coherency prediction for the IEEE 39-bus system, the proposed method is more than 3.78 and 2.39% superior to [17] and [19], respectively. This is despite the fact the proposed method does not require any post-fault data, which substantially reduces the prediction time.

To better exemplify the effects of DEMs on generator grouping, the distribution of 328 training samples associated with faults occurring at line 2–3 of the IEEE 39-bus system is shown in Fig. 6. This figure shows that even two DEMs can form a proper visual distinction between different coherent groups. With TDSs of multiple DEMs, the proposed framework generates enough input features to enable the ensemble DT to reach high prediction accuracies.

To measure the performance of the proposed technique for the system topological changes, detailed prediction results are reported in Table IV for different layouts. This table shows the prediction accuracy for the original network ($N-0$) is a little bit (<1%) more than the other situations. The main reason behind this behavior is that almost 80% of the dataset is related to the normal topology and thus the ML-technique can develop a better prediction model. It is empirically seen in simulations that the proposed method can fairly

TABLE VI
EFFECTS OF PMU NOISE ON THE AVERAGE PREDICTION ACCURACY (%)

| Item | Test situation | 39-bus | | 145-bus | |
|---|---|---|---|---|---|
| | | Without noise | With Noise | Without noise | With Noise |
| Stability status prediction | [17] with PFC = 60 | 96.96 | 96.33 | 95.47 | 94.70 |
| | Proposed method | **98.35** | **98.26** | **98.11** | **97.97** |
| Coherency prediction | [17] with PFC = 60 | 92.67 | 92.11 | 92.32 | 91.69 |
| | Proposed method | **96.45** | **96.29** | **95.17** | **94.88** |

TABLE VII
DETAILED COMPUTATIONAL TIME FOR DIFFERENT IEEE TEST SYSTEMS

| Item | Details | | 39-Bus | 145-Bus |
|---|---|---|---|---|
| Offline training (time in minutes) | Run and record TDS for all scenarios | | 77.446 | 210.741 |
| | Form DEMs, run and record TDS for DEMs | | 128.632 | 257.916 |
| | Extract features and find $\Omega_{\boldsymbol{AB},i}^{RT}$ | | 9.055 | 53.197 |
| | Build prediction models | | 14.726 | 106.933 |
| | Total | | 229.859 | 628.787 |
| Online application for a single test sample (average time in seconds) | Steady-state and during fault | Form DEMs | 0.259 | 0.418 |
| | After fault clearing | Run TDS for DEMs | 0.216 | 0.201 |
| | | Extract features | 0.013 | 0.011 |
| | | Apply prediction model | 0.008 | 0.008 |
| | | Total | 0.237 | 0.220 |

predict different topological changes if the coherency group is available in the dataset.

Aimed at evaluating the effects of different parameters involved in the generator aggregation process on the final prediction accuracy, a sensitivity analysis is carried out. In this test, the equivalent inertia constant, damping constant, synchronous reactance, etc. are calculated in detail based on (9)–(11) and the effects of modeling error on some parameters are investigated with the results reproduced in Table IV. To this end, $\boldsymbol{k}_a(i)$, $\boldsymbol{k}_b(i)$, and $\delta_i$ are assumed to be the same for all generators of a coherent group, which results in an average value for each parameter. Table IV shows the excitation system reflects the most effects on prediction accuracies and the subtransient time constant represents the least. The solutions obtained reveal that the aggregation technique, introduced in Section III, suitably performed the equivalent modeling and was well integrated into the proposed framework.

To determine the effects of PMU noise on the proposed technique, IEEE 39- and 145-bus systems are considered in additional examination. White Gaussian noise with a signal-to-noise ratio equal to 34 dB is added to the dataset, i.e., to both training and test data, and the training process is repeated. The results obtained for the proposed method are illustrated in Table VI and compared to those of [17] with PFC=60. This table shows that, on average, stability status prediction accuracy decreases by 0.70% and 0.12% for [17] and the proposed method, respectively; in other words, the proposed method is 5.83 times less sensitive to PMU noise. Because the developed method only relies on the last steady-state snapshot of the grid, PMU noise may minimally change the stability status of the system. The proposed method mistakenly classified the stability status in only 10 and 13 cases for the IEEE 39- and 145-bus systems, respectively. The simulations empirically show that the operating point of those cases was close to the stability boundaries, and thus the inclusion of PMU noise affected the decision made.

To assess the curse of dimensionality, detailed computational time of the proposed method for two IEEE test cases is reproduced in Table VII. This table shows the offline training process took about 230 minutes in the case of the IEEE 39-bus network, of which 137.69 minutes is spent to run parallel TDSs for DEMs and to extract features. Simply put, with respect to the existing stability prediction methods where direct PMU data are employed, the proposed technique increases the computational burden of the offline model training stage by 60%. However, the simulation time still meets the engineering requirements; in the IEEE 145-bus network, the training process takes around 10.5 hours. In addition, the dataset generation can be conducted in parallel using high-performance computing (HPC), which can substantially reduce the training time for larger networks.

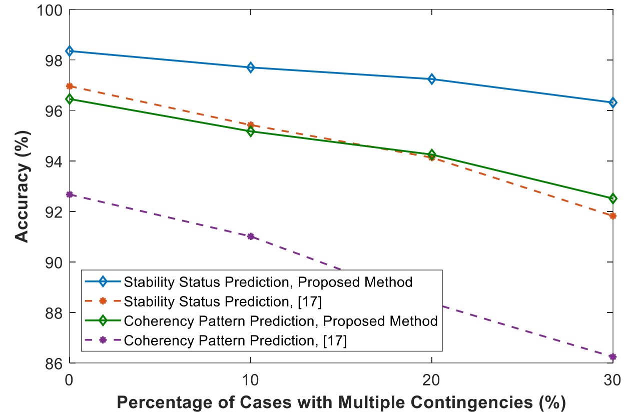

Fig. 7. Performance of different methods on test data for the IEEE 39-bus system considering multiple contingencies.

Table VII shows that, in online applications, DEMs are formed in the steady-state situation and the nodal aggregation is conducted during a fault. Once the fault clears, parallel TDSs are triggered and the features are extracted. Simulations empirically show that running 20 s of post-fault simulations using the PSSE transient stability module takes about 8 ms for an equivalent two-machine network. With this, and in the case of the IEEE 39-bus system, the overall post-fault simulation time of the proposed method is 0.237 s. Comparing this value with about 1 s waiting time of [17] and [19] for $PFC = 60$, the proposed framework is almost 4.2 times faster in online applications.

### B. *Second Scenario*

In this scenario, the effects of multiple contingencies and load model uncertainties on the performance of the proposed method are evaluated. Moreover, to better reflect the contrast with the existing techniques, the solutions obtained are compared with [17] (while $PFC = 60$).

First, one-fifth of the database associate with the IEEE 39-bus system is randomly selected and called “test data” from now on. A portion of the test cases (10, 20, or 30%) is randomly selected. Then, an extra three-phase fault with a random location and random fault duration (2–15 cycles) is applied to each case. This extra contingency is set to be cleared by tripping the faulted line, and the main fault is applied at least 5 to 60 cycles (selected by chance) after clearance of the first contingency. A new TDS is carried out for each case; because $N-1$ and $N-2$ contingencies were considered while generating offline scenarios, the multiple faults did not lead to an unforeseen coherency pattern. The prediction models trained in the first scenario are applied to all test cases and the results obtained are depicted in Fig. 7. This figure shows the proposed stability prediction method is robust against multiple contingencies; in the worst case, the accuracy only drops by 2.07% with respect to the first scenario. Applying [17], which directly uses PMU data, to the same dataset, the prediction accuracy reduction is more than 5.14%, i.e., the pro-

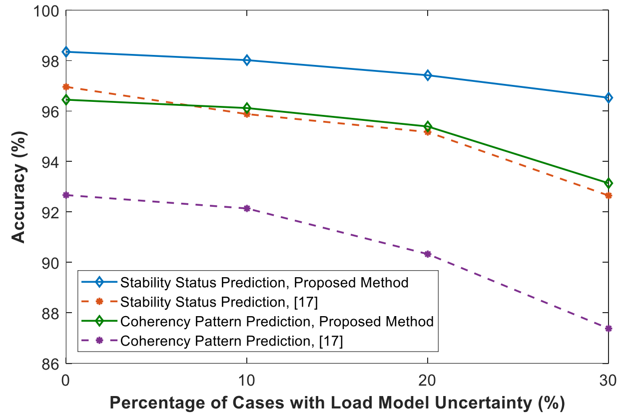


Fig. 8. Performance of different methods on test data for the IEEE 39-bus system considering load model uncertainties.

TABLE VIII
PERCENTAGE OF CASES IN WHICH STABILITY STATUS OR COHERENCY PATTERN ARE CHANGED WITH THE INCLUSION OF UNCERTAINTY IN THE LOAD MODEL ON TEST DATA OF IEEE 39-BUS SYSTEM (%)

| Cases in which | Percentage of loads with uncertain model | | |
|---|---|---|---|
| | 10% | 20% | 30% |
| Stability status changed | 0.31 | 1.06 | 1.93 |
| Coherency pattern changed | 0.42 | 1.27 | 2.69 |

posed method is almost 2.48 times more robust than [17] when it comes to multiple contingencies.

Aimed at evaluating the effects of load model uncertainties, a portion of loads (10, 20, or 30%) in each test case is randomly selected. Then, the electrical model of each chosen load is randomly modified so that up to 30% of the base demand is assumed to be modeled by CP and up to 30% by IM; the remaining portion is modeled by a CI load model. A new TDS is conducted for each test case and the data obtained are recorded. While no new coherency pattern appeared in this test, load model uncertainties did cause some changes in the stability status or coherency pattern of test cases compared to Table II. Such changes are reproduced in Table VIII, which shows that, in the worst case, the stability status of 1.93% of cases and the coherency pattern of 2.69% of the test cases changed. Hence, the sensitivity of generator dynamic behavior to load model uncertainty is considered to be 2.69% in this study.

To assess the sensitivity of the prediction techniques to load model uncertainties, the offline prediction models trained in the first scenario are applied to the new test data. The data obtained are plotted in Fig. 8, which shows that, in the worst case, the accuracy of the proposed stability status prediction is reduced by 1.82% with respect to the first scenario, which is still 1.37 (i.e. (4.31-1.82)/1.82) times better than [17]. Because the proposed framework is founded on a simulation-based classification approach in which all load models are considered as CI during both offline and online applications, it is unable to track the effects of load model uncertainties and any changes in the stability status or coherency pattern of the generators. However, because it does not directly use PMU data, it has more robust performance against load model uncertainties compared to the available techniques [5]–[20].

## VI. CONCLUSION

This paper proposed a novel solution framework for prediction of generator dynamic behavior. Relying on simulation-based classification, an aggregation method was employed to derive DEMs for each network. Parallel TDSs were carried out for equivalent two-machine DEMs, which resulted in a set of features used for training the prediction models. A conditional mutual information-based approach was used for feature selection and an ensemble DT was implemented to build up prediction models. The proposed framework was successfully applied to several IEEE test systems and the results obtained were compared to existing techniques. The findings and comparisons demonstrated the effectiveness of the proposed method in real-time applications.

Further research may be conducted to enhance the prediction performance of the proposed method by adopting other machine learning techniques. In addition, the proposed offline training procedure can be improved with a framework that can handle significant changes in the operating condition of the system in a timely manner suitable for day-ahead operation.

## REFERENCES


[1] E. Vaahedi, *Practical Power System Operation*, Piscataway, NJ: Wiley, 2014.

[2] C.W. Lin, J.S. Thorp, "A novel method to compute the closest unstable equilibrium point for transient stability region estimate in power systems," *IEEE Transactions on Circuits and Syst.*, vol. 44, no. 7, pp. 630–635, 1997.

[3] W. Li, M. Wang, and J.H. Chow, "Real-time event identification through low-dimensional subspace characterization of high-dimensional synchrophasor Data," *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 4937–4947, 2018.

[4] A.S. Dobakhshari, "Wide-area fault location of transmission lines by hybrid synchronized/unsynchronized voltage measurements," *IEEE Trans. Smart Grid*, vol. 9, no. 3, pp. 1869–1877, 2018.

[5] S. Dasgupta, M. Paramasivam, U. Vaidya, and V. Ajjarapu, "PMU-based model-free approach for real-time rotor angle monitoring," *IEEE Trans. Power Syst.*, vol. 30, no. 5, pp. 2818–2819, 2015.

[6] H.N.V. Pico and B.B. Johnson, "Transient stability assessment of multi-machine multi-converter power systems," *IEEE Trans. Power Syst.*, vol. pp, no. 99, pp. 1–11, 2019.

[7] P. Ju, H. Li, C. gan. Y. Liu, Y. Yu, and Y. Liu "Analytical assessment for transient stability under stochastic continuous disturbances," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 2004–2014, 2018.

[8] A. Gupta, G. Gurrala, and P.S. Sastry, "An online power system stability monitoring system using convolutional neural networks," *IEEE Trans. Power Syst.*, vol. 34, no. 2, pp. 868–872, 2019.

[9] J.J.Q. Yu, D.J. Hill, A.Y.S. Lam, J. Gu, and V.O.K. Li, "Intelligent time-adaptive transient stability assessment system," *IEEE Trans. Power Syst.*, vol. pp, no. 99, pp. 1–10, 2017.

[10] R. Zhang, Y. Xu, Z.Y. Dong, and K.P. Wong, "Post-disturbance transient stability assessment of power systems by a self-adaptive intelligent system*," IET Gener. Transm. Dis.*, vol. 19, no. 5, pp. 296–305, 2015.

[11] S.M. Mazhari, N. Safari, C.Y. Chung, and I. Kamwa, "A hybrid fault cluster and Thévenin equivalent based framework for rotor angle stability prediction," *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 5594–9603, 2018.

[12] H. Supreme, L.A. Dessaint, I. Kamwa, and A. Heniche-Oussedik, "Development of new predictors based on the concept of center of power for transient and dynamic instability detection," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 3605–3615, 2018.

[13] H. Mehrjerdi, S. Lefebvre, M. Saad, and D. Asber, "A decentralized control of partitioned power networks for voltage regulation and prevention Against disturbance propagation," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1461–1469, 2013.

[14] M.R. Salimian and M.R. Aghamohammadi, "Intelligent out of step predictor for inter area oscillations using speed-acceleration criterion as a time matching for controlled islanding," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 2488–2497, 2018.

[15] M.H.R. Koochi1, S. Esmaeili, and R. Fadaeinedjad, "New phasor-based approach for online and fast prediction of generators grouping using decision tree," *IET Gener. Transm. Dis.*, vol. 11, no. 6, pp. 1566–1574, 2017.

[16] M. Rashidi and E. Farjah, "LEs based framework for transient instability prediction and mitigation using PMU data," *IET Gener. Transm. Dis.*, vol. 10, no. 14, pp. 3431–3440, 2016.

[17] T. Guo and J.V. Milanovic, "Online identification of power system dynamic signature using PMU measurements and data mining," *IEEE Trans. Power Syst.*, vol. 31, no. 3, pp. 1760–1768, 2016.

[18] P.N. Papadopoulos, T. Guo, and J.V. Milanovic, "Probabilistic framework for online identification of dynamic behavior of power systems with renewable generation," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 45–53, 2018.

[19] S.M. Mazhari, N. Safari, C.Y. Chung, and I. Kamwa, "A quantile regression-based approach for online probabilistic prediction of unstable groups of coherent generators in power systems," *IEEE Trans. Power Syst.*, vol. 99, no. 99, pp. 1–10, 2019.

[20] S.A. Siddiqui, K. Verma, K.R. Niazi, and M. Fozdar, “Real-time monitoring of post-fault scenario for determining generator coherency and transient stability through ANN,” *IEEE Trans. Indus. App.*, vol. 54, no. 1, pp. 685–692, 2018.

[21] Y. Song, W. Wang, Z. Zhang, H. Qi, and Y. Liu, “Multiple event detection and recognition for large-scale power systems through cluster-based sparse coding,” *IEEE Trans. Power Syst.*, vol. 23, no. 6, pp. 4199–4209, 2017.

[22] J. Bialek, et al., “Benchmarking and validation of cascading failure analysis tools,” *IEEE Trans. Power Syst.*, vol. 31, no. 6, pp. 4887–4900, 2016.

[23] L. Korunovic, J.V. Milanovic, S.Z. Djokic, K. Yamashita, S.M. Villanueva, and S. Sterpu, “Recommended parameter values and ranges of most frequently used static load models,” *IEEE Trans. Power Syst.*, vol. 33, no. 3, pp. 5923–5934, 2018.

[24] M. Gordon, “Impact of load behavior on transient stability and power transfer limitations,” *IEEE PES General Meeting*, pp. 1–8, 2009

[25] H. Elahi, “Transient stability of power systems with non-linear load models using individual machine energy functions,” Ph.D. Dissertation, Iowa State University, 1983.

[26] M.L. Ourari, L.A. Dessaint, and V.Q. Do, “Dynamic equivalent modeling of large power systems using structure preservation technique,” *IEEE Trans. Power Syst.*, vol. 21, no. 3, pp. 1284–1295, 2006.

[27] R. Nath and S. S. Lamba, “Development of coherency-based time domain equivalent model using structure constraints,” *Proc. Inst. Elect. Eng.*, vol. 133, no. 4, pt. C, pp. 165–175, May 1986.

[28] P. Goos and D. Meintrup, *Statistics with JMP: Graphs, Descriptive Statistics and Probability*, NJ: Wiley, 2015.

[29] J. Novovicova, P. Somol, M. Haindl, and P. Pudil, “Conditional mutual information based feature selection for classification task,” *Springer Lecture Notes in Computer Science*, vol. 4756, pp. 417–426, 2007.

[30] [online] https://siemens.force.com/SEMC2/s/article/PSS-E-User-Support