Received June 21, 2021, accepted July 18, 2021, date of publication July 26, 2021, date of current version August 2, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3099459

Time Series-Analysis Based Engineering of High-Dimensional Wide-Area Stability Indices for Machine Learning

RAOULT TEUKAM DABOU[1], (Student Member, IEEE), INNOCENT KAMWA[1], (Fellow, IEEE), C. Y. CHUNG[2], (Fellow, IEEE), AND CHUMA FRANCIS MUGOMBOZI[3], (Member, IEEE)

[1]Department of Electrical and Computer Science Engineering, Laval University, Quebec, QC G1V 0A6, Canada
[2]Department of Electrical and Computer Engineering, University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada
[3]Department of Power Systems Simulation and Evolution, Research Institute of Hydro-Quebec (IREQ), Varennes, QC J3X 1S1, Canada

Corresponding author: Raoult Teukam Dabou (raoult.teukam-dabou.1@ulaval.ca)

This work was supported by the Power Electronics and Industrial Control Laboratory (LEEPCI), Department of Electrical and Computer Engineering, Laval University, through the Discovery Research Program of Canada National Sciences and Engineering Research Council (NSERC).

ABSTRACT Information representative of actual power system dynamics is usually buried in masses of phasor measurement unit (PMU) data. To take full advantage of these data in early anticipation of stability loss, we propose to implement the high dimensional stability index (HDSI). This method allows the extraction of more than 500-labeled attributes describing generator response signals, such as speed and rate of change of transient energy function (RoCoTE). A combined 31 functions are computed from spectrum analysis based on the Periodogram and Welch methods, Lyapunov exponents, and wavelet transform approaches. The test databases are built by simulating faults on each line in the IEEE 39- and 68-bus networks. Applying comparative time-series analysis to such signal responses to disturbances then highlights the texture matrix of the stability attributes. A 10-fold support vector machine (SVM) is used to implement a HDSI-based stability prediction model, with its performance then compared to the artificial neural network (ANN), decision trees (DT), random forest (RF), and adaptive boosting (AdaBoost) models available in the statistical package R. While most methods performed similarly, with ∼100% accuracy on test cases using the same set of HDSI-based attributes, the RF classifier with its associated Gini feature importance allows for explicit feature ranking and interpretation, which results in prioritization of frequency-domain over time-domain features.

INDEX TERMS Stability signal responses, time-series classification, stability attributes, wide-area severity indices, fast Fourier transform, Welch method, periodogram, Lyapunov exponent, wavelet, machine learning, dynamic state estimation, dynamic security assessment.

NOMENCLATURE

$x_n, \dot{x}$	State and State Derivative variable
(x_0, u_0)	Point of equilibrium of state and input
y, u_r	Output and input vectors
γ, t	Cut-off stability value and time
$\overline{\Delta\delta_i}, \psi_i$	Maximum deviation and stability index
$\Delta u_i, \Delta x_i$	Small deviation of the input and state variable
δ_i, W_{di}	Rotor angle and Rate of change of transient energy (RoCoTE)
$\lambda_i(x)$	Lyapunov exponent
$\Lambda(x)$	Matrix coefficient
T'_{d0i}, T'_{q0i}	d & q transient open circuit time constant
Ω^c, N_d	Set of disturbances and Number of databases
f_i, g_i	System state and output functions
N_l, N_c, N_g	Number of lines, contingencies and generators
N_f	Number of HDSI based features in data mining
x_d, x_q	Direct (d) & quadrature (q) reactance
$PC-1, PC-2$	First and second principal component

The associate editor coordinating the review of this manuscript and approving it for publication was Hiu Yung Wong.

e'_d, e'_q	d & q axis transient electromotive voltage
Ψ, ω	Wavelet transform and frequency variable
x'_d, x'_q	d & q axis transient reactance
x_w, R	Segment samples and window size in periodogram and Welch methods
m, k	Number of blocks and available frames in periodogram and Welch methods
SVM	Support vector machine
DT	Decision trees
RF	Random forest
ANN	Artificial neural network
AdaBoost	Adaptive boosting

I. INTRODUCTION

Time-series of post-disturbance response signals contain dormant and non-transparent information that is useful for understanding the dynamics of power systems in real time [1]. The analysis and rapid interpretation of the physical meaning of these data usually requires considerable experience in electrical engineering and signal processing [2]. It thus suitable to provide artificial intelligence (AI) experts and junior engineers with a simple method to extract and classify transparent and easily interpretable stability predictor attributes. Highlighting the physical meaning of these attributes allows movement away from the black-box approach of dynamic stability evaluation. The difficulty in accessing real data from power networks motivated [3] to propose a fast and simple algorithm for generating stability data for different configurations of IEEE test systems in Simscape Power System (SPS).

With advances in deep learning applied to time-series [4], several authors have turned to this approach to assess network stability by analyzing post-fault signals [5]. Although classifications by deep learning leads to relatively satisfactory accuracy, the raw signal-based decision engine is opaque and difficult to explain to experts responsible for network planning and operation. As pointed out in [6], this opacity is an important limitation to the maintenance of such models and their adoption by the industry. In this context, the characterization of temporal responses prior to machine learning by physically interpretable attributes, via notions such as the maximum of spectral energy density or maximum of Lyapunov exponents, becomes an interesting alternative. However, until now, time-series classification based on transparent attributes has been limited to a small number of possible attributes (generally from 1 to 10 according to authors), with notable difficulty in generalizing the attributes thus chosen to new networks and new configurations, as some of the proposed attributes are a function of the inertia of the machines and topology of the network [7]. This paper proposes to overcome these limitations by adopting a highly comparative time-series approach (HCTSA) [8] to extract a massive number of attributes describing the network response signals to perturbations, without making assumptions about the network topology and parameters. The central tenet of the proposed approach is that by increasing drastically the type and number of physically interpretable "PMU data based catastrophic indicators" [17] involved in the machine learning (i.e. moving from a low dimensional to high-dimensional feature space), we should be able to improve performance and transparency of the power system stability prediction model. The goal is therefore not to outperform any machine learning or deep-learning algorithms, as no new such algorithms is proposed in this paper, but rather to extending considerably the set of candidate features with a physical meaning in order to improve the performance of existing machine learning tools, while allowing physical interpretations or reformulation of the resulting data-driven stability prediction models into transparent fuzzy logic rules. In fact, HDSI will simultaneously enhance classification performance of existing machine learning and increase the transparency and readability of the underlying algorithms decision rules.

Along similar lines, the work of [9] presents a method to analyze the dynamic safety of disturbed power systems by a machine learning approach that uses voltage phasors to build a neural network model for predicting the status (stable/unstable) of the post-fault power system. The maximum Lyapunov exponent is proposed in [10] to predict the rotor angle stability in real time from phasor measurement unit (PMU) data. The authors argue that, during operation, the generators remain synchronous if the linear system associated to the Lyapunov coefficients is asymptotically stable. In [11], classification rules for the rapid assessment of post-disturbance stability are developed from decision trees (DT) and fuzzy rule approaches. The authors extract and classify stability attributes from the Hydro-Quebec database of PMU signal responses recorded after short and long duration disturbances under changing topologies. In [12], the focus is on the importance of the time-frequency distribution of power system electromechanical energy oscillations in network stability after a disturbance. Short-time Fourier transform, fuzzy logic, and NN approaches are used in [13] to extract and classify stability predictor attributes. In [14], an algorithm based on discrete Fourier transform and interpolation is proposed to assess the dynamics of generators in real-time using the field-programmable gate array. DT applied to the integral square of generator angle and integral square of bus angle indices is used in [15] to detect events and predict loss of synchronism.

In light of the above literature, the present paper focuses on the combined use of Lyapunov exponents, wavelet and Fourier transforms, and spectral energy using periodograms and Welch methods to extract stability precursor attributes from time-series of generators rotor speeds and RoCoTE derived from Dynamic State Estimation (DSE) utilizing PMU data [1]. The power system dynamics information from the Lyapunov exponents allow the extraction of attributes related to stability or instability by quantifying the eigenvalue variations and sensitivity degree of the underlying

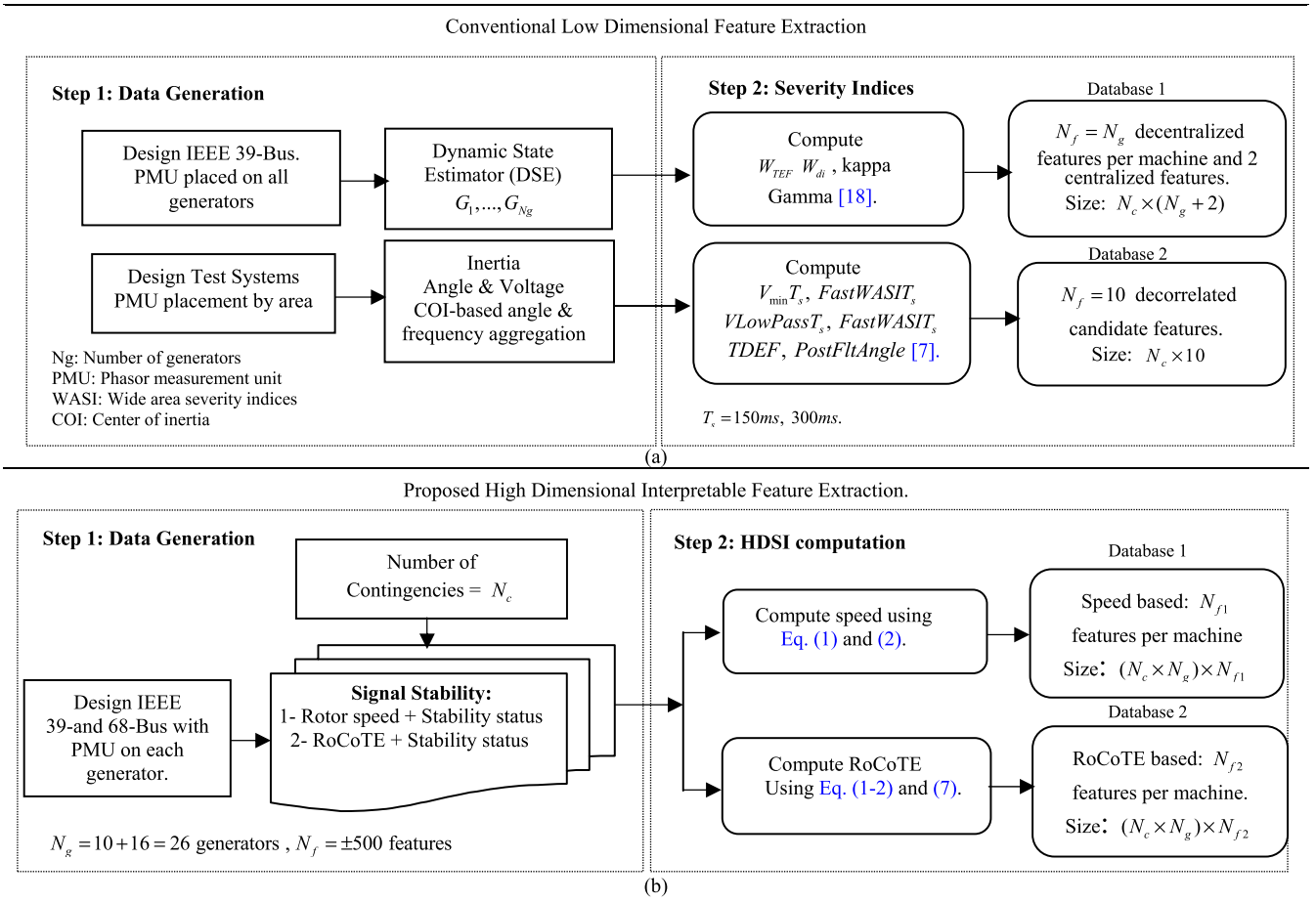

FIGURE 1. **Block diagram of conventional (a) and proposed HDSI (b) Feature extraction scheme.**

linearized system. Similarly, the spectral analysis of generator response signals captures disrupting energy processes driving the system towards the brink of instability. It is implemented by applying rectangular and Hamming windows, resulting in the periodogram and Welch methods. A level 2, 3, 4, 5 time-frequency decompositions, carried out on the Daubechies and Symlet wavelets, facilitates the statistical extraction of the energy attributes contained in stability signals when the latter are highly nonstationary. These methods enable the definition of a cluster matrix of attributes characterizing the power system dynamic stability status in post-disturbance stage.

Fig. 1 summarizes the differences between the conventional and proposed approaches for feature extraction to enhance the effectiveness and transparency of the data mining-based stability prediction. In one conventional approach, Fig. 1. (a), the stability signals recorded at the machine levels are used to calculate for each machine the RoCoTE and Transient Energy Function (TEF) features. Likewise, the concept of wide-area severity indices is used in another conventional approach [7] to assess the stability status based on ensemble decision trees fed with 10 features characterizing the Hydro-Québec network in post-disturbance conditions. The second conventional approach performs instability detection based on a single decentralized indicator per machine combined with two centralized indicators computed from global response of the machines [18]. Both methods can be used to analyze/decide wide-area dynamic stability with less than ten physically interpretable features, see step 2, Fig. 1. (a). Although the conventional approach relies on a small set of features useful for understanding the dynamics of power systems in real time, it does not explore the full spectrum of candidate features which is potentially very large, and therefore, it does not allow for a hierarchization of the complementary features within the set (i.e. importance ranking of time-domain vs frequency domain attributes), which is required for an in-depth physical understanding of the predictor ability to capture stability phenomena.

By contrast, the new paradigm of HDSI based features extraction, Fig. 1. (b), consists in applying to any available time-response signal arising from short-time monitoring, five complementary signal processing engines simultaneously.

Considering two basic signals responses, namely, the rotor speed and RoCoTE, we can derive from the basic HDSI concept, two databases each representing a different data

mining paradigm. The implementation of Eq. 5 and Eq. (8-15) corresponds to speed signal-based HDSI, which is the basic HDSI in this paper due to its used of an unprocessed raw physical signal.

The implementation of Eq. (5-7) and Eq. (8-15) corresponds to yet another variant of HDSI method using decentralized RoCoTE, which is a byproduct of the decentralized DSE. Each of the two databases, obtained by implementing Eq. (5) to (15), can generate 500+ attributes per single generator rotor speed or RoCoTE signals respectively, through HCTSA. This study compares the performances of each HDSI method on IEEE 39- and 68-bus test systems for which source code is available on request.

The proposed data mining approach is first built in Matlab using the 10-fold SVM approach whose performance is next compared to stability predictions from the following learning algorithms in R: artificial neural network (ANN), random forest (RF), adaptive boosting (AdaBoost), and decision trees (DT) [25]. The two first principal components (PC-1 and PC-2) are also used to classify the stable and unstable attributes on a linear projection of a hyperplane to help visualizing and understanding the stability phenomena. Furthermore, using the Gini importance of attributes produced by the random forests to rank the ∼500 candidate features, the energy-related attributes are found to be more discriminative of the dynamic stability status than Lyapunov exponents while the speed-based HDSI are deemed more effective than RoCoTE based HDSI. These are few major findings from the proposed HCTSA-based understanding, interpretation, and classification of stability time-series responses.

The rest of the paper is organized into five sections. Section II describes the proposed high-dimensional stability indices. Section III presents the simulation results of high-dimensional stability features extraction and classification based on the IEEE 39-, 68- and 39 + 68 -bus systems. Our methodology considers a single database consisting of $N_l \times N_c \times N_g$ time-series from these two widely different systems to obtain more general conclusions with respect to the physical interpretation and ranking of stability attributes. According to [34], the former test system has mostly transient stability issues while the second is dominated by small-signal stability issue, and therefore, combining the two networks in a single data base has the potential provides two complementary views of the stability phenomena in a single prediction model. Section IV presents a comparison between the basic HDSI using speed signal and the RoCoTE HDSI variant across several well-known machine learning engines built in Matlab and R. Section V presents the conclusions of the work and future perspectives.

II. PROPOSED HIGH-DIMENSIONAL STABILITY INDICES

The analysis and interpretation of a power system during its operation requires the resolution of electromechanical equations describing the relationships between all generators.

A. DYNAMIC STATE ESTIMATION

The differential/algebraic equations (DAE) characteristic of power system dynamics is defined by [16]:

$$\begin{aligned} \dot{x}_i &= f_i\left(x_1, x_2, \ldots, x_n; u_1, u_2, \ldots, u_r; t\right)(a) \\ 0 &= g_i\left(x_1, x_2, \ldots, x_n; u_1, u_2, \ldots, u_r; t\right)(b) \end{aligned} \tag{1}$$

where the first set (1a) represents the differential equations and the second set (1b) represents the algebraic equations which are typically associated to the network constraints in steady-state at fundamental frequency. With some compatible initial conditions (x_0, y_0), i.e.: $0 = g\left(x_0, y_0\right); f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n,\ g : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m$, equations (1) can be rewritten as:

$$\begin{aligned} \dot{x}_i &= f_i\left[(x_0 + \Delta x), (u_0 + \Delta u)\right] \\ 0 &= g_i\left[(x_0 + \Delta x), (u_0 + \Delta u)\right] \end{aligned} \tag{2}$$

$$\begin{aligned} \Delta\dot{x}_i = &\frac{\partial f_i}{\partial x_1}\Delta x_1 + \ldots + \frac{\partial f_i}{\partial x_n}\Delta x_n + \frac{\partial f_i}{\partial u_1}\Delta u_1 \\ &+ \ldots + \frac{\partial f_i}{\partial u_n}\Delta u_n \end{aligned} \tag{3}$$

Similarly, the output variables Δy_i are expressed as functions of the state x_i and input variables u_i by:

$$\Delta \mathrm{y}_i = \frac{\partial g_i}{\partial x_1}\Delta x_1 + \ldots + \frac{\partial g_i}{\partial x_n}\Delta x_n + \frac{\partial g_i}{\partial u_1}\Delta u_1 + \ldots + \frac{\partial g_i}{\partial u_n}\Delta u_n \tag{4}$$

In the presence of disturbances, the linearization of Eq. 2 makes it possible to characterize the stability of oscillation modes of electrical networks using the Lyapunov theory. Most software (Matlab/Simulink, PSSE, PSLF, PowerWorld, DSATools, etc.) implements Eq. (1-4) to perform an offline analysis of the power system dynamics [1]. After fault clearance, signals from the synchronous generators (directly measured or extrapolated from direct measurements) are used to calculate decentralized stability indicators [18]. In real-time conditions, PMU data are used in DSE to derive the internal dynamic state of the generators from terminal measurements for the purpose of online stability status assessment [17].

B. INSTABILITY DETECTION IN MULTIMACHINE SIMULATION

The rotor angle $\delta_i^k(t)$ of generator k for contingency $i \in \Omega^c$ is used to compute the maximum deviation $\overline{\Delta\delta}_i(t)$ between any pair $k, m \in \Omega^g$ of generators at instant t and compare it to a cut-off value $\gamma = \pi$ [34].

$$\psi_i(t) = \frac{\gamma - \overline{\Delta\delta}_i(t)}{\gamma + \overline{\Delta\delta}_i(t)} = \begin{cases} 1 & \text{Stable } if\ \psi_t \succ 0 \\ 0 & \text{Unstable } if\ \psi_t \leq 0 \end{cases} \tag{5}$$

$$\overline{\Delta\delta}_i(t) = \max\left(\left|\delta_i^k(t) - \delta_i^m(t)\right|\right) \tag{6}$$

When a disturbance affects a power system, we have two possible status: if $\overline{\Delta\delta}_i(t)$ is high, the generators will have a tendency of falling apart and the difference between γ (which is taken equal to π) will be small and hence the indicator will be close to zero [33]. Thus, at each simulation step, a binary decision on the stability status of the sample is calculated

and simultaneously stored in the database. During operation of the power grid, the decentralized stability indicator called the RoCoTE function (W_{di}) is calculated using DSE for generator i according to [18]:

$$W_{di} = \frac{dW_i}{dt} = -\left[\frac{T'_{d0i}}{(x_d - x'_d)}\left(\frac{de'_{qi}}{dt}\right)^2 + \frac{T'_{q0i}}{(x_q - x'_q)}\left(\frac{de'_{di}}{dt}\right)^2\right] \quad (7)$$

The collection of rotor speed signals, along with Eq. 5 and Eq. 7, are used to extract the discriminative attributes of the power system dynamics through application of HCTSA. These stability status (“0” or “1”) is used to label the training time-series snapshots (rotor speed and RoCoTE), i.e. a binary number 1 or 0 is assigned as output class or categorial variable if the signals (rotor speed and RoCoTE) correspond to an instable or stable contingency respectively.

C. POWER ENGINEERING FEATURES EXTRACTION

The stability database consisting of rotor speeds and RoCoTE is obtained by simulating three-phase ground fault on each line and close to generator bars of IEEE 39 –and 68 –bus. These sample simulations are chosen to highlight the power system dynamics when faced with line tripping, generator loss, load switch on and load switch off. The faults being applied at $t = 1s$, are eliminated after two cycles (i.e. post-fault time: $1 + 2/60$ sec) for IEEE 39-bus and after 12 cycles (i.e. post-fault time: $1 + 12/60$ sec) for IEEE 68-bus, we have essentially allocated a maximum of $250ms$ data window for decision making after fault clearing. Table 1 presents the configuration of disturbance scenario simulated in the IEEE 39- and 68-bus test systems.

It is well known that the performance of the learning machine is not only related to the learning algorithm but also to the quality of the input data with respect to specific tasks. The block diagram shown in Fig. 2 summarizes to steps for extracting such features using the characteristics of the Lyapunov exponents, statistics of power spectrum, and wavelet transform applied to the database of time-series signal responses. Thus, from the selected four Lyapunov functions, five spectral analysis functions, and 22 wavelet transformations, more than 500 stability attributes are extracted per signal.

The Daubechies and Symlet wavelet transforms applied to the stability database facilitates the extraction of discriminant information in the frequency domain. The wavelet transform is a set of various transforms that uses functions located in both real space and real time to describe the scaling properties of the stability signal. The continuous wavelet transform expression, calculated from the mother wavelet Ψ, is defined in [19] by:

$$\langle x, \Psi_{a,b}\rangle = \frac{1}{\sqrt{a}}\int_t x(t)\,\Psi\left(\frac{t-b}{a}\right)dt \quad (8)$$

where: a and b represent dilation and translation factors.

TABLE 1. Fault simulation scenario of the IEEE 39 - and - 68 bus.

Disturbance		Fault Location & configuration
Three-Phase Ground (3-Ph G)	Load-on Events	Fault on each generator bars, Load change: +20% MW Gain
	Load-off Events	Fault on each generator bars, Load change: -20% MW Loss
	Line Tripping	Each line: Applied at 50% (IEEE 39 - bus) and 75% (IEEE 68 -bus).
	Generator Loss	Fault on each generator bars, Generator change: -10% & -30% MVA Loss

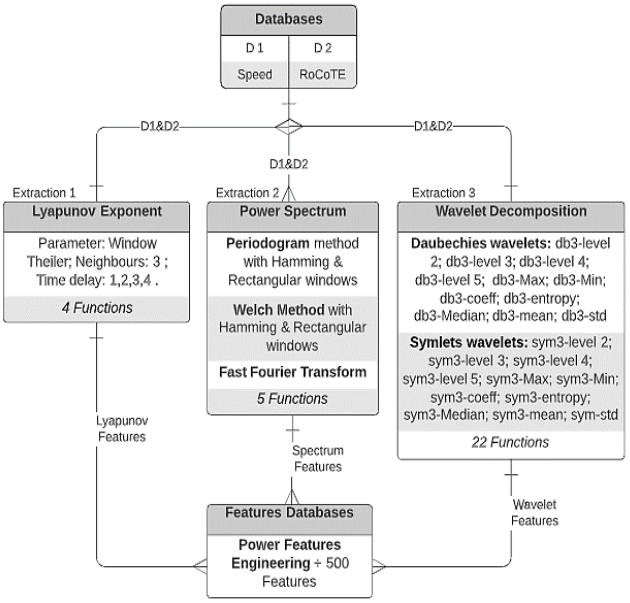

FIGURE 2. Block diagram of HDSI-based features extraction.

These spectral features evolve over time, thus allowing to identify common time-varying patterns in rotor speed and RoCoTE signals, and perform accurate time-localization of event based on low and high frequency band filters in (8). The wavelet decomposition allows therefore to detecting change point introduced in the power system by fault occurrence. The periodogram, Welch method, and fast Fourier transforms permit estimation of the spectral density of each time-series over short-time windows following the fault clearing. The periodogram method divides the stability feature in several segments of precise length that are progressively shifted and process with fast Fourier transform, $X(f)$ define as follows:

$$\mathrm{FFT}(X(f)) \triangleq \int_t x(t)e^{-j2\pi ft}dt \quad (9)$$

The spectrum amplitude of the stability signals from (9) using the fast Fourier transform applied to post-fault time-series data enables the energy introduced by the disturbance into the power grid to be highlighted. The periodogram $P_{f,M}(\omega)$ is defined as the squared magnitude of the discrete

time Fourier transform of x divided by M, as per [20]:

$$x_w(n) = w(n)x(n) \tag{10}$$

$$\mathrm{P}_{f,M}(\omega) = \frac{1}{M}\left|\sum_{n=0}^{M-1} x_w(n)e^{-j\omega n}\right|^2 \tag{11}$$

The implementation of Welch $\mathrm{W}_{f,M}(\omega)$ method amounts to applying several periodograms using a sliding window [21]:

$$x_m(n) = w(n)x(n+mR) \tag{12}$$

$$\mathrm{W}_{f,M}(\omega) = \frac{1}{M}\left|\sum_{n=0}^{M-1} x_m(n)e^{-j2\pi nk/N}\right|^2 \tag{13}$$

When a disturbance is applied to the power system, it is subject to slow or brief variations in initial conditions. The Lyapunov exponents derived from these variations can allow the prediction of stability or instability during the post-disturbance period, based on the trajectories of generator variables (speed, voltage, RoCoTE, etc.). The largest Lyapunov exponent is considered a suitable metric for stability prediction. Referring to the dynamic system of (Eq. 14) with $x \in X \subset \mathbb{R}^N$ and a solution $\Phi(t, x)$, we define the limit matrix using Λ:

$$\Lambda(x) = \lim_{t\to\infty}\left[\frac{\partial\Phi(t,x)^T}{\partial x}\frac{\partial\Phi(t,x)}{\partial x}\right]^{1/2t} \tag{14}$$

Let $\Lambda_i(x)$ be the eigenvalues of limiting matrix $\Lambda(x)$ [22]. The Lyapunov exponents $\lambda_i(x)$ are defined as:

$$\lambda_i(\mathrm{x}) = \log\Lambda_i(\mathrm{x}) \tag{15}$$

The negative (positive) value of the maximum Lyapunov exponent implies the power system is stable (unstable) around a nearby equilibrium point. The characteristic nonlinear differential equations of generator dynamics are used to determine the power system equilibrium. Then, the simulation of each fault generates precise exponents that can be positive (network is unstable) or negative (network is stable), according to the electrical network sensitivity and initial conditions.

D. K-FOLD SVM BASED FEATURES CLASSIFIER

From the concepts described above in Eq. (8-15), 31 functions (see Fig. 2) with various parameterizations are applied to each time-series of rotor speeds and RoCoTE to extract discriminative decentralized attributes associated with each generator. Because the support vector machine (SVM) is one of the most powerful classifiers in terms of the tradeoff between effectiveness and complexity [26], especially in the context of power stability prediction [31], it was selected for implementing the basic HDSI in Matlab [8]. Its use in the predictive analysis of dynamic stability allows the signals in the database to be separated into two classes: stable or unstable. To enhance its generalization capability, the proposed SVM algorithm uses K-fold cross-validation [23], [24] to ensure that each available time-series is used in both the training and validation sets. This approach minimizes the expected overall error in the full database by selecting the hyperplane that best discriminates the stability signals to maximize the distance margin between it and the nearest sample. Given a set of models $f(x, \alpha)$ indexed by tuning parameter α, denote by $f^{-k}(x, \alpha)$ the α^{th} model fit with the k^{th} part of the data removed. Then, for this set of models, we define:

$$CV(\hat{f}, \alpha) = \frac{1}{N}\sum_{i=1}^{N} L\left(y_i, \hat{f}^{-k}(x_i, \alpha)\right) \tag{16}$$

where $CV\left(\hat{f}, \alpha\right)$ provides an estimate of the test error curve, and we find the tuning parameter $\hat{\alpha}$ that minimizes it. Our final model is $f\left(x, \hat{\alpha}\right)$, which we then fit to all data. In other words, for K-fold learning, the signals are randomly partitioned into $K = 10$ groups. Each in turn, one of the groups is used as a test set for a training model and the other four groups are used as a training set at each iteration. Therefore, the training attribute boundaries are used to define the plausible location of the test attributes in relation to the splitting hyperplane.

TABLE 2. Confusion matrix configuration.

Actual Class	Predicted Class	
	Stable	Unstable
stable	C_{11}	C_{01}
Unstable	C_{10}	C_{00}

E. PERFORMANCE MEASURES FROM THE CONFUSION MATRIX

The confusion matrix in Table 2, also known as an error matrix, is used to evaluate and visualize the performance of each stability classifier. It consists of three metrics providing a detailed overview of the generalization performance of the training model on the test database.

The accuracy metric in (17) defines the ratio of correctly predicted stability status (stable/unstable) to the total number of training data.

$$Accuracy(\%) = \frac{C_{11} + C_{00}}{C_{11} + C_{00} + C_{10} + C_{01}} \tag{17}$$

The reliability metric in Eq. 18, allows to evaluating the ability of the classifier to correctly predicting the instabilities of the power systems.

$$\mathrm{Re}liability\,(\%) = \frac{C_{00} - C_{01}}{C_{00}} \tag{18}$$

The security metric in Eq. 19 allows evaluating the capacity of the classifier to correctly predicting the stable status of the power systems:

$$\mathrm{Sec}urity(\%) = \frac{C_{11} - C_{10}}{C_{11}} \tag{19}$$

TABLE 3. Performance metrics of machine learning of stability prediction using rotor speed and RoCoTE based HDSI.

Test Systems	Software	Learning Algorithms	Signal processing	Accuracy (%)	Reliability (%)	Security (%)
IEEE 39 -bus	Matlab	10- fold SVM	**Speed**	**100**	**100**	**100**
			RoCoTE	99.45	99.06	100
	Rattle	Random Forest	**Speed**	**99.59**	**99.76**	**99.35**
			RoCoTE	99.86	99.76	100
		AdaBoost	**Speed**	**99.72**	**100**	**99.35**
			RoCoTE	99.86	99.76	100
		Artificial Neural Network	**Speed**	**100**	**100**	**100**
			RoCoTE	99.72	99.53	100
		Decision Trees	**Speed**	**98.91**	**99.06**	**98.69**
			RoCoTE	98.64	98.34	99.02
IEEE 68 -bus	Matlab	10- fold SVM	**Speed**	**99.66**	**99.50**	**99.76**
			RoCoTE	92.78	91.77	92.51
	Rattle	Random Forest	**Speed**	**99.90**	**99.75**	**100**
			RoCoTE	98.60	97.61	99.20
		AdaBoost	**Speed**	**99.85**	**99.75**	**99.92**
			RoCoTE	98.50	97.35	99.20
		Artificial Neural Network	**Speed**	**99.75**	**99.63**	**99.84**
			RoCoTE	97.83	97.22	98.14
		Decision Trees	**Speed**	**98.22**	**96.57**	**99.20**
			RoCoTE	92.98	84.90	96.81
IEEE 39 - and 68-bus	Matlab	10- fold SVM	**Speed**	99.68	97.86	99.90
			RoCoTE	97.77	78.16	99.65
	Rattle	Random Forest	**Speed**	99.96	99.55	100
			RoCoTE	99.78	97.93	99.96
		AdaBoost	**Speed**	99.96	99.55	100
			RoCoTE	99.81	98.81	99.94
		Artificial Neural Network	**Speed**	99.92	99.92	99.94
			RoCoTE	99.33	95.21	99.72
		Decision Trees	**Speed**	98.46	83.64	99.71
			RoCoTE	97.41	60.70	99.96

III. SIMULATION RESULTS

The power system dynamics are studied through two databases generated by simulating line tripping and three-phase to ground fault on each line and at generator terminal busses of the IEEE 39- and 68-bus networks. The partial load switching on/off (±20% MW) and generation shedding (−10% & −30% MVA) allows to enhancing the richness of both training databases. Instead of applying the HDSI method on the two systems separately, we instead combine the two sets of data in the hope of finding a single, more general, predictor. This idea was first proposed in [7], where it was pointed out that learning the physically inspired "stability concept" on networks with widely different dynamic stability attributes (first swing vs. multi-swing, oscillatory vs. voltage stability, etc.) will result in a more general predictor encapsulating a robust stability definition which retains significance across various dynamic phenomena. By combining the two datasets, with the first dominated by transient stability phenomena while the second is dominated by small-signal stability type of events [34], it is possible to derive a single machine-learning based stability predictor, able to generalize the stability features to all types of network dynamics; this leads to a single model that can also work potentially for never seen networks or novel operating conditions on previously seen networks. Interestingly, Table 3 confirms that combining the two test systems in a single stability database results in a RoCoTE based DT with only a 60% reliability, which compares to 98% and 84% for the 39-bus and 68-bus test systems respectively, when they are taken alone. By contrast, the RoCoTE based random forest model results in a greater than 97.6% reliability rate for all three scenarios which demonstrates that the random forest model has enhanced generalization capabilities, which allows it to predict the stability phenomena across a wider range possible of systems dynamics, provided that the stability database is rich enough.

A. TIME SERIES FOR STABILITY PREDICTION

The IEEE 39- and 68-bus networks respectively consist of 34 and 66 lines and 10 and 16 generators. The simulations of the networks presented in Fig. 3. (a) and Fig. 3. (b) were performed in SPS. The consist of 740 and 2080 time-series respectively, assuming only one signal response (such as rotor speed) is recorded per generator. In addition, for each sample recorded, a stability status ("0" or = "1") is defined during the simulation according to Eq. (5-6).

Note that considering the RoCoTE as a second generator signal doubles the number of stability time-series in each database (i.e. 5 640 time-series) which was investigated indeed, but without reporting the results here for reasons explained in the Discussion section. The configuration of the database allows each time-series to be named based on

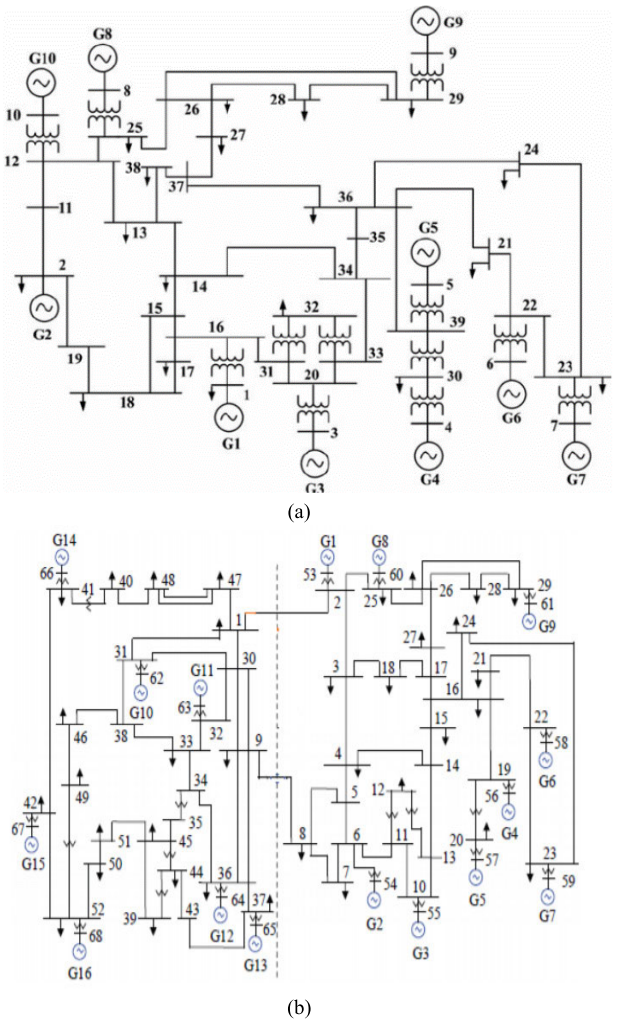

FIGURE 3. Single line diagram of the (a) IEEE 39-bus and (b) IEEE 68-bus systems.

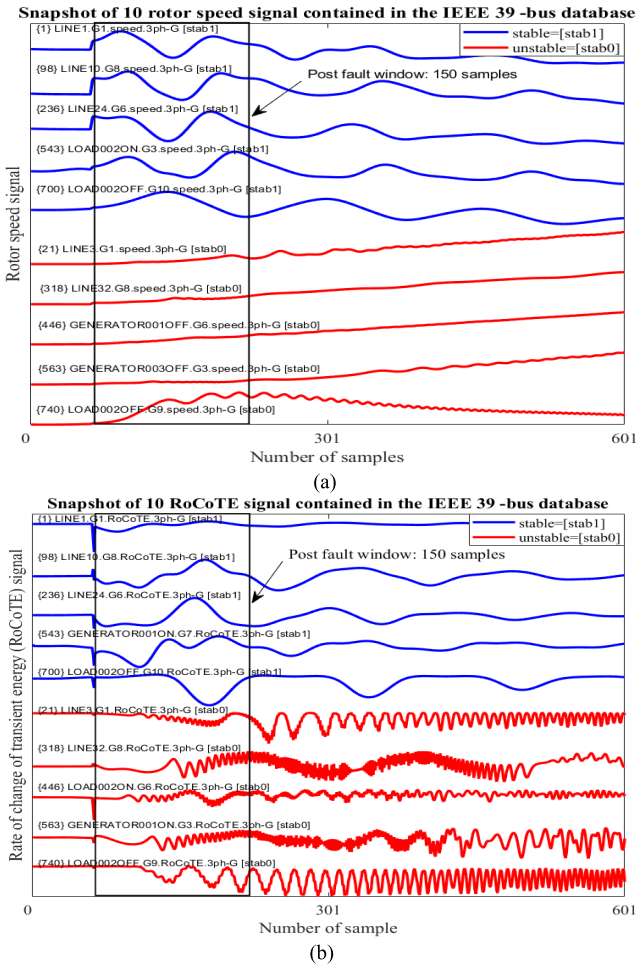

FIGURE 4. Time-series: (a) rotor speed and (b) rate of change transient energy (IEEE 39 -bus).

information about the ID number of the generator that produced the recorded speed signal, the line on which the fault was applied, the type of fault, and the global stability status of the network in post-disturbance dynamics. To illustrate the database built by simulating the two power systems in Fig. 3. (a) and Fig. 3. (b), we present sample snapshots of rotor speed trajectories in Fig. 4. (a) and RoCoTE in Fig. 4. (b), with stable cases (blue) and unstable cases (red).

Each pre-processed stability signal is tagged with six labels that provide the information on signal ID in database, fault location, ID generator, signal response, type of fault and stability status. These figures show that the RoCoTE present more fluctuations than the rotor speed signal, which may be expected from the derivative operation. However, they also display only minor visual differences between times-series from stable and unstable cases in the early stage of post-disturbance dynamics, over 150 samples, which underscores the challenge of predicting the stability from time-series data directly.

B. POWER ENGINEERING FEATURES MAPPING

To take full advantage of the data and improve the visibility of the dynamic stability features following HCTSA, clustering of rows and columns of the databases is performed. This hierarchical reorganization between rows and columns enables the calculation of the row distance and column distance metrics, according to the database stability label. The sigmoid function is used to normalize the stability database and display the attribute weights from lowest (blue) to highest (red). The online signals are no longer represented by many colors but only by blue (stable) or red (unstable). The rows of the attribute matrix represent the number of signals present in the database whereas the columns represent the number of attributes calculated for each signal. The attribute matrix derived from the rotor speed alone is shown in Fig. 5.

C. DYNAMIC STABILITY FEATURE CLASSIFICATION

A preliminary comparison of the statistical distribution of stability attributes extracted from databases is carried out in order to inspect visually and highlight their ability to discriminate the dynamics of the power system. Thus, for each calculated attribute, it is possible to know what percentage of

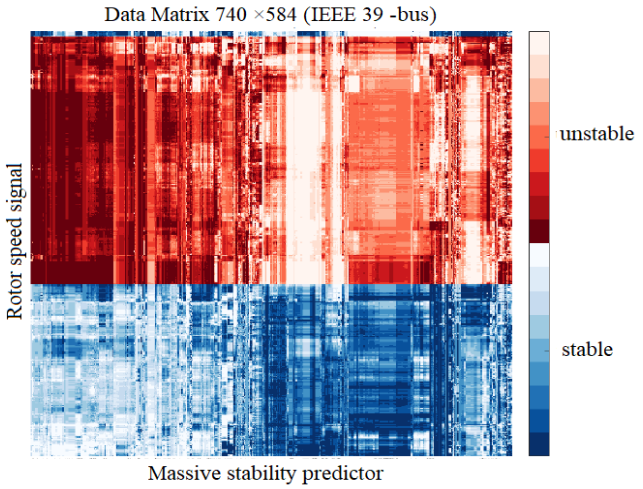

FIGURE 5. Attributes data matrix for rotor speed (IEEE 39 -bus).

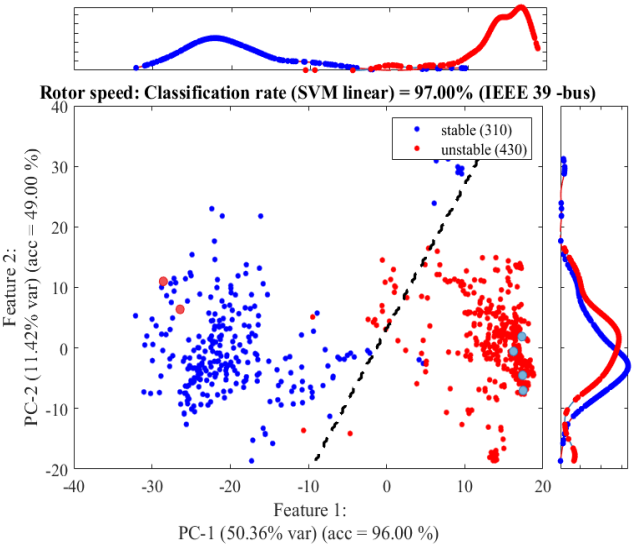

FIGURE 6. Linear SVM classification of rotor speed (IEEE 39-bus).

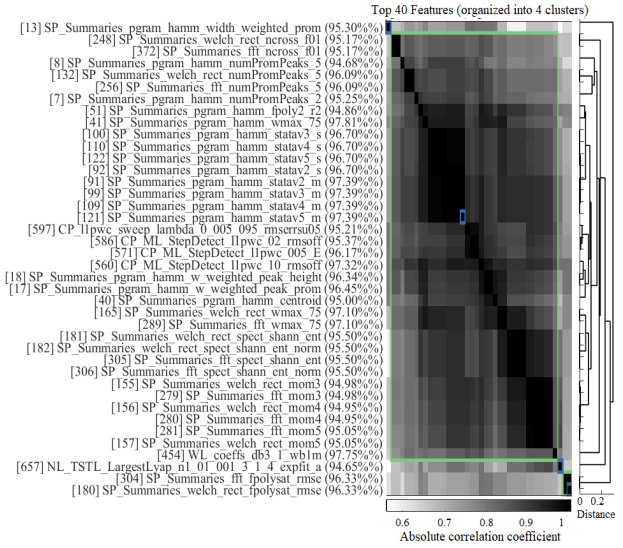

FIGURE 7. Clustering top feature of rotor speed (IEEE 39- bus).

the sample data it classifies correctly as stable and unstable cases according to the predefined labels.

By observing the rotor speed database, we find in Fig. 6, that PC-1 and PC-2 are enough to separate the plane approximately by a line, with most of the blues and reds being on either side of the line. To improve the separation, it would take a hyperplane of at least 10 dimensions (number of significant principal components according to Rattle), the distribution in blue and red confirm that PC-1 has a very negative value when unstable while PC-2 has a pronounced overlap between blue and red which gives limited precision to separate stable and unstable cases: the accuracy in predicting correctly the stability of a case using PC-1 only is 96% vs 49% for PC-2 only. A 50% accuracy means that the classification is not more accurate that a toss-up game, confirming that the overlap in distributions makes the separation between stable/unstable very poor. The patterns in Fig. 6 are the PC-1 (high) and PC-2 (low) distributions as a function of stability (blue = stable, red = unstable). The SVM algorithm maximized the smallest existing signed distance between the stable and unstable signals [32], see Fig. 6.

While Fig. 2 presents the rationale behind the full set of potential features (up to +500 numbers derived from 31 basic functions), the 40 most significant attributes of the rotor speed database (as obtained by minimizing redundancy and correlation between features) are grouped into four clusters, as shown in Fig. 7. The Daubechies wavelet (up to 97.75%) and fast Fourier transform (up to 97.81%) best discriminate the rotor speeds compared to Lyapunov exponents (94.65%). This result, which highlights the relevance of frequency domain features as the best wide-area severity indices [17], will be further confirmed in a latter section using the more rigorous random forest Gini importance of features [28].

D. HDSI-BASED PREDICTOR PERFORMANCE

The implementation of 10-fold cross validation SVM based on rotor speed and RoCoTE data alone results in the confusion matrix shown in Fig. 8. The rotor speed classifier correctly detects 310 stable signals (over 310) and 430 unstable signals (over 430), see Fig.8. (a). Whereas, the RoCoTE classifier correctly detects 310 stable signals (over 310) and 426 unstable signals (over 430), see Fig.8. (b).

Similarly, the implementation of 10-fold cross validation SVM based on the rotor speed and RoCoTE data results of IEEE 68 –bus. The rotor speed classifier correctly detects 1261 stable signals (over 1264) and 812 unstable signals (over 816), see Fig.9. (a). The RoCoTE classifier correctly detects 1176 stable signals (over 1264) and 754 unstable signals (over 816), see Fig.9. (b).

While comparable to existing literature, these figures are inferior to the 10-fold SVM performance using rotor speed signals (Fig. 8). We can conclude that the rotor speed achieves better performance compared to the RoCoTE based on the same HCTSA process with a 100% vs 99.5% accuracy, see Fig. 8 (a)-(b).

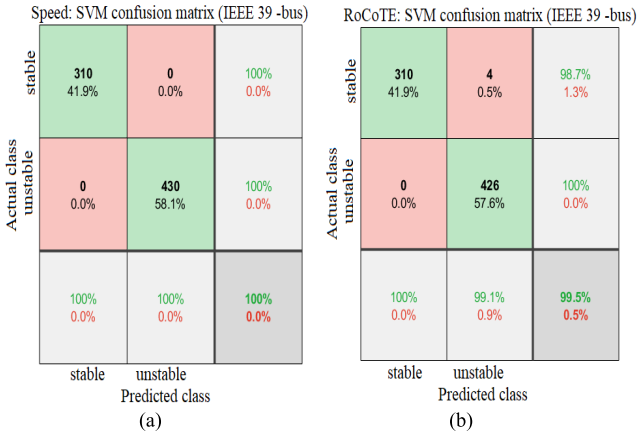

FIGURE 8. 10-fold SVM confusion matrix for (a) rotor speed and (b) RoCoTE (IEEE 39-bus).

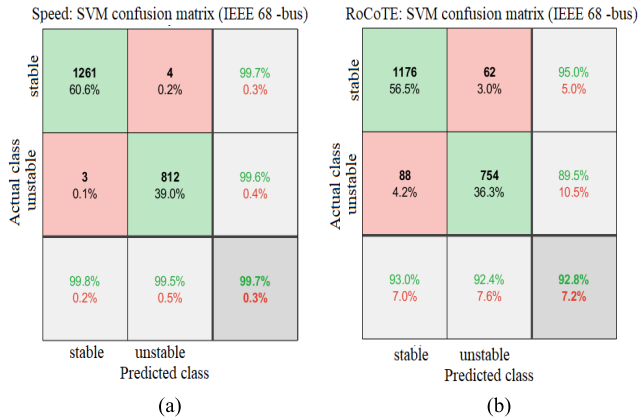

FIGURE 9. 10-fold SVM confusion matrix for (a) rotor speed and (b) RoCoTE (IEEE 68-bus).

E. COMPARISON WITH EXISTING MACHINE LEARNING TOOLS

The data-mining tool, Rattle for R, is used to evaluate the performance of the HDSI on stability data [25]. This requires preliminary preparation of the stability data into a format that Rattle can understand [26], [27]. The file associated with both the IEEE 39- and 68-bus networks contains 500+ stability attributes tags (columns) and 2 820 signals (rows). HDSI's SVM 10-fold linear classifier model built from databases of rotor speeds and the RoCoTE show good accuracy, reliability and security compared to the decision trees method. For the same databases, the Rattle AdaBoost, ANN and RF classifiers are all highly accurate, secure and reliable, and their performance is slightly better than that of the 10-fold SVM classifier, Table 3. Overall, the HDSI features from rotor speed data result in very few errors in the classification of stability attributes.

However, with the same HCTSA attributes fed to the mainstream machine learning algorithms RF, AdaBoost, and ANN, they all result in nearly the same good performance, which makes the choice of which to use a matter of preference or availability. The main idea of this work, highlighted in Fig. 2, was to demonstrate that the application

TABLE 4. CPU time (in sec.) of HSDI.

N° of busses	HDSI	10-fold SVM	RF	Boost	ANN	DT
39	325.17	5.28	9.80	12.95	55.31	7.45
68	880.57	14.04	70.36	34.82	113.09	13.50
39+68	7738.67	460.01	494.77	54.28	120.94	29.67

of several attribute extraction methods (Lyapunov exponent, power spectrum and wavelet decomposition) on the raw rotor speed database can perform well, with little intervention of the AI expert thanks to HCTSA.

The results in Table 3 confirms that this is indeed true and further allows us to affirm that 1) the classifier is a transparent machine learning stability predictor with explainable input features and 2) its performance depends more on learning attributes than the classification algorithm. Table 4 presents in its first column the computational load of HDSI estimation for the test systems considered in this paper. The performance was assessed on a DELL computer configured with the Intel processor i7-7700HQ 4-core running at 2.80 GHz with 16 GB of RAM. Although the CPU time is relatively significant, the actual code is in Matlab scripting language and therefore, it can be made faster using C-programming.

In addition, since each time series (or subset of time-series) is analyzed individually using HCTSA, the problem of computing HDSI is naturally parallelizable, by splitting for instance, the 10 machines found in the 39-bus systems into 10 parallel tasks. For completeness, Table 4 also presents in the five last columns, the respective computational requirements of the 5 machine learning algorithms studied in this paper. It appears that the Boost technique offers the best compromise between classification accuracy and computational burden while the random forest and 10-fold SVM CPU times grow exponentially with the problem size.

IV. DISCUSSION

An open question that remains is which attributes among the ∼500 is the most relevant in terms of discriminative efficacy. We have improved the quality of Fig. 10, thus highlighting feature importance in terms of Gini impurity reduction, according to the random forest machine learning method [28].

These results are very consistent with the clustering approach in Fig. 7, showing that frequency domain attributes are the most efficient decentralized severity indices [17]. Indeed, the first Lyapunov-based coefficient is ranked 61. To highlight the critical importance of this finding in building efficient and understandable classifiers, Fig. 11 presents box plots Fig. 11. (a) and distributions Fig. 11. (b) of one of the three-frequency domain variables compared to the best Lyapunov exponent.

Whereas, the latter features overlapping distribution of stable and unstable cases, superposition of the two Welch spectra results in a clear-cut histogram showing extreme values correspond to unstable cases in both the IEEE 39 -bus. The performance of HDSI derived from rotor speed

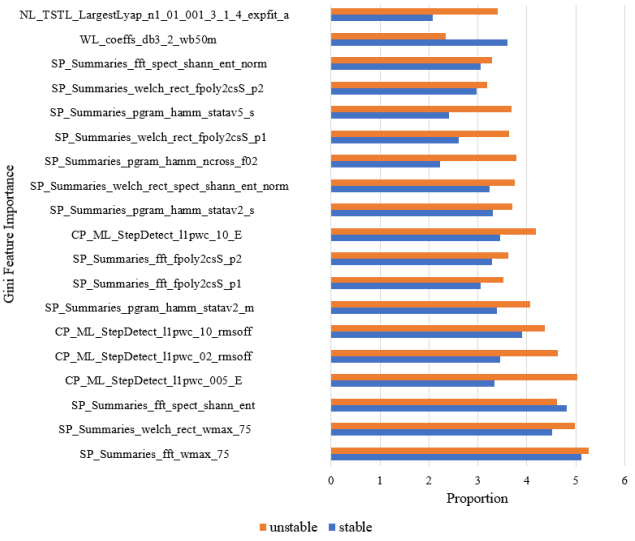

FIGURE 10. Rotor speed feature importance in terms of Gini impurity reduction (IEEE 39 -bus).

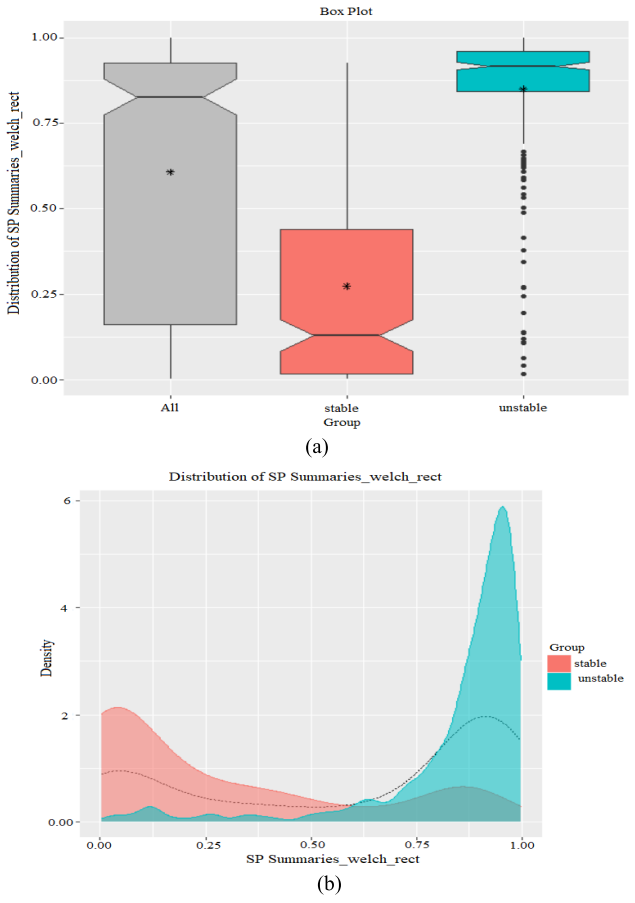

FIGURE 11. Characterizing distributions of top frequency domain attribute: (a) box plot, (b) density distribution (IEEE 39 + 68 -bus).

alone is so good using either 10-fold SVM, AdaBoost, or RF predictors that little is gained on the existing IEEE 39-plus 68-bus database by stacking HDSI from rotor speed and RoCoTE. In fact, it was found in a sharp contrast to Fig. 6, that a principal components decomposition of the HDSI based attributes associated to RoCoTE performed rather poorly in splitting stable and unstable cases in two crisp classes using a linear SVM (Fig. 12).

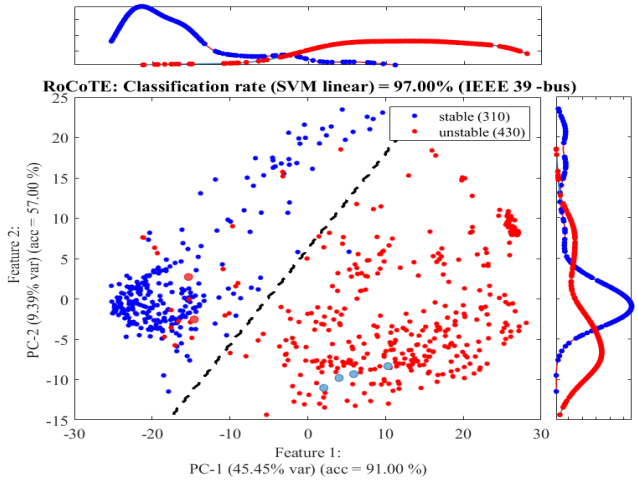

FIGURE 12. Linear SVM classification of RoCoTE (IEEE 39 -bus).

Only two attributes (PC-1 & PC-2) out of 500 are used for the SVM classification of the stability database, see Fig. 6 & Fig. 12.

While for the sake of simplicity it is possible, even preferable, to analyze dynamic stability with a few predictors (less than 10) as shown in [7] and [18], these attributes are usually selected in an ad-hoc manner with limited justification and without any comparison with the myriad of alternative possibilities existing in the literature (e.g. in time-domain vs frequency domain). These methods using a single feature such as the maximum angle in the COI reference, the peak of the rotor speed spectral density or the maximum Lyapunov exponent do not compare and rank these attributes in order to define their intrinsic predictive power, leaving the reader puzzled by the question of which of the basic features is actually the most informative about stability phenomena. The proposed HDSI is an attempt to reconciliate all the main time-series based features found in the existing literature in a unified framework, thus enabling their fair comparison and ranking while providing the machine learning with much more degrees-of-freedom to explore during the classifier model building. Admittedly, with the consideration of +500 physically explainable attributes describing each generator signal response, we agree that this approach which expands the search space from a dozen of attributes to hundreds of attributes may seem on the surface like an increased complexity. But this perception missed two aspects of the method stressed in this paper:

1) The computation of the +500 attributes is highly automated, numerically efficient, and masks a lot of digital signal processing complexity to the user

2) The only way to select the best physically sound features (in absence of prior knowledge or expertise) is to explore in the same framework, all categories of candidates reported

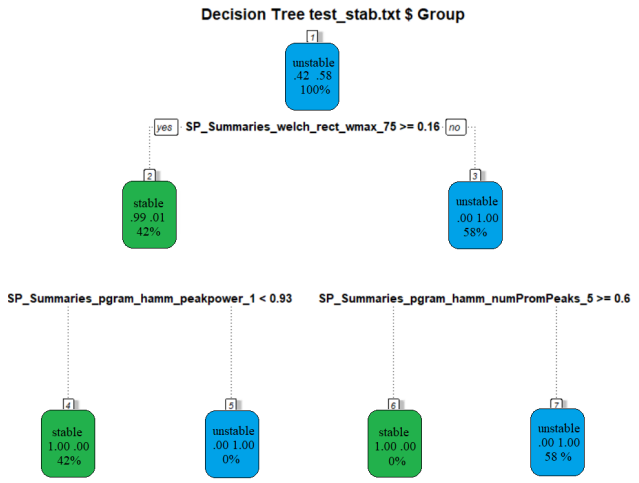

FIGURE 13. Decision trees classification of rotor speed (IEEE 39- bus).

by experts in existing literature, which means to computing Fourier, Welch, Wavelet, Lyapunov based features simultaneously to enable comparisons.

3) Machine learning tools such as ensemble decision trees are designed to naturally navigate through many features to derive models with use the most informative features only. However, it is the duty of the data scientist to provide the machine learning the full set of informative and uncorrelated features to work with, which can only happen in the rare circumstances where the data scientist is also a domain expert.

Therefore, the HDSI approach allows to build a large set of possible features which enable the data scientists who are non-domain experts to define hierarchically the predictive power of the attributes and to give them a physical interpretation as in Fig. 6, according to the fundamental dynamics of the power systems. For illustration, despite the high number of inputs, the decision tree in Fig. 13 is quite simple if not trivial to understand since it uses only 3 variables among the +500-labeled to achieve instability prediction with a 99% reliability on the IEEE 39-bus test system providing a strong evidence of the discriminatory power of the three features used in building the DT. These critical attributes turned out to be all a subset of the frequency domain attributes illustrated in Fig. 10 and they could not have been so easily discovered in ad-hoc manner without exploring the full set of +500 attributes and ranking their predictive power. Such a simple tree can be easily converted into a transparent and robust fuzzy rule classifier [29], see Fig. 13.

However, comparing the performances of the speed and RoCoTE based HDSI in systematic way, and prioritizing the frequency domain features with a greater confidence would require the simultaneous addition of another test system, such as the IEEE 50-machine test case, or consideration of the voltage stability phenomenon in addition to transient and oscillatory stability conditions. In the future, we also plan to compare the performance of HDSI based transparent machine learning of stability status with direct opaque time-series based deep learning using standard algorithms available in Matlab [30] together with larger test systems and larger data sets.

V. CONCLUSION

Existing methods for attribute analysis and extraction usually require significant expertise for the physical interpretation of signals. This work aims to develop a promising high dimensional stability index, called HDSI, for wide-area stability monitoring and control. This new tool can extract and classify more than 500 attributes describing each generator response signal obtained from dynamic state estimation based on PMU data. This is an important alternative for the promotion of machine learning applications to power systems dynamics. To extract the stability attributes, 31 different analysis functions from fast Fourier and wavelet transforms, periodogram and Welch spectral methods and Lyapunov exponents are applied to each post-disturbance response signal recorded during fault simulation on each of the lines and close to generator bars of the test networks. Such attributes are then adjacently correlated and standardized to improve visibility and interpretation of stability features hidden in the signal responses. The most discriminating attributes are clustered according to their percentages and then classified into two stability categories using the 10-fold support vector machine.

The proposed HDSI-based features are used in prediction models built in Matlab and Rattle using 10-fold SVM, DT, RF, ANN, and AdaBoost learning algorithms. All models performed well, except for the standard DT model that has unsatisfactory reliability. The two main findings arising from using the HDSI method are as follows: 1) the top attributes in terms of discriminative power and encapsulating the stability phenomenon are frequency domain-based features and 2) 10-fold SVM, RF, ANN, and AdaBoost learning algorithms result in similar performance when fed with the same HDSI-based attributes from highly discriminative time-series analysis. The latter conclusion enables the development of more transparent stability predictors built using DSE based synchronous machines response signals in post-disturbance conditions.

REFERENCES

[1] J. Zhao, J. Qi, Z. Huang, A. P. S. Meliopoulos, A. Gomez-Exposito, M. Netto, L. Mili, A. Abur, V. Terzija, I. Kamwa, B. Pal, and A. K. Singh, "Power system dynamic state estimation: Motivations, definitions, methodologies, and future work," *IEEE Trans. Power Syst.*, vol. 34, no. 4, pp. 3188–3198, Jul. 2019.

[2] B. D. Fulcher, M. A. Little, and N. S. Jones, "Highly comparative time-series analysis: The empirical structure of time series and their methods," *J. Roy. Soc. Interface*, vol. 10, no. 83, Jun. 2013, Art. no. 20130048.

[3] R. T. Dabou and I. Kamwa, "Rapid design method for generating power system stability databases in SPS for machine learning," in *Proc. IEEE Can. Conf. Electr. Comput. Eng. (CCECE)*, Aug. 2020, pp. 1–6.

[4] H. I. Fawaz, G. Forestier, J. Weber, L. Idoumghar, and P.-A. Müller, "Deep learning for time series classification: A review," *Data Mining Knowl. Discovery*, vol. 33, no. 4, pp. 917–963, Jul. 2019.

[5] S. K. Azman, Y. J. Isbeih, M. S. E. Moursi, and K. Elbassioni, "A unified online deep learning prediction model for small signal and transient stability," *IEEE Trans. Power Syst.*, vol. 35, no. 6, pp. 4585–4598, Nov. 2020.

[6] I. Kamwa, S. R. Samantaray, and G. Joós, "On the accuracy versus transparency trade-off of data-mining models for fast-response PMU-based catastrophe predictors," *IEEE Trans. Smart Grid*, vol. 3, no. 1, pp. 152–161, Mar. 2012.

[7] I. Kamwa, S. R. Samantaray, and G. Joos, "Catastrophe predictors from ensemble decision-tree learning of wide-area severity indices," *IEEE Trans. Smart Grid*, vol. 1, no. 2, pp. 144–158, Sep. 2010.

[8] B. D. Fulcher and N. S. Jones, "*Hctsa*: A computational framework for automated time-series phenotyping using massive feature extraction," *Cell Syst.*, vol. 5, no. 5, pp. 527–531, 2017.

[9] T. Niimura, H. S. Ko, H. Xu, A. Moshref, and K. Morison, "Machine learning approach to power system dynamic security analysis," in *Proc. Power Syst. Conf. Expo.*, Oct. 2004, pp. 1084–1088.

[10] J. Yan, C.-C. Liu, and U. Vaidya, "PMU-based monitoring of rotor angle dynamics," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 2125–2133, Nov. 2011.

[11] I. Kamwa, S. R. Samantaray, and G. Joos, "Development of rule-based classifiers for rapid stability assessment of wide-area post-disturbance records," *IEEE Trans. Power Syst.*, vol. 24, no. 1, pp. 258–270, Feb. 2009.

[12] D. R. Ostojic, "Spectral monitoring of power system dynamic performances," *IEEE Trans. Power Syst.*, vol. 8, no. 2, pp. 445–451, May 1993.

[13] I. Kamwa, R. Grondin, and L. Loud, "Time-varying contingency screening for dynamic security assessment using intelligent-systems techniques," *IEEE Trans. Power Syst.*, vol. 16, no. 3, pp. 526–536, Aug. 2001.

[14] R. Yadav, A. K. Pradhan, and I. Kamwa, "A spectrum similarity approach for identifying coherency change patterns in power system due to variability in renewable generation," *IEEE Trans. Power Syst.*, vol. 34, no. 5, pp. 3769–3779, Sep. 2019.

[15] S. M. Rovnyak, D. W. Longbottom, D. C. Vasquez, and M. N. Nilchi, "Response-based event detection for one shot wide-area stability controls," *Monitoring and Control using Synchrophasors in Power Systems with Renewables*, I. Kamwa, C. Lu, L. Zhu, Eds. London, U.K.: The Institution of Engineering and Technology Press, Jul. 2020, ch. 8, pp. 1–24.

[16] P. Kundur, *Power System Stability and Control*. New York, NY, USA: McGraw-Hill, 1994.

[17] G. Ravikumar and S. A. Khaparde, "Taxonomy of PMU data based catastrophic indicators for power system stability assessment," *IEEE Syst. J.*, vol. 12, no. 1, pp. 452–464, Mar. 2018.

[18] A. Paul, I. Kamwa, and G. Joos, "PMU signals responses-based RAS for instability mitigation through On–The fly identification and shedding of the run-away generators," *IEEE Trans. Power Syst.*, vol. 35, no. 3, pp. 1707–1717, May 2020.

[19] L. Wei and H. Put, "Application of wavelet transformation for voltage stability in distributed power system network," in *Proc. Chin. Control Decis. Conf.*, Yantai, Shandong, Jul. 2008, pp. 116–119.

[20] S. A. Fulop and K. Fitz, "Algorithms for computing the time-corrected instantaneous frequency (reassigned) spectrogram, with applications," *J. Acoust. Soc. Amer.*, vol. 119, no. 1, pp. 360–371, Jan. 2006.

[21] K. Barbe, R. Pintelon, and J. Schoukens, "Welch method revisited: Nonparametric power spectrum estimation via circular overlap," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 553–565, Feb. 2010.

[22] S. Dasgupta, M. Paramasivam, U. Vaidya, and V. Ajjarapu, "Real-time monitoring of short-term voltage stability using PMU data," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 3702–3711, Nov. 2013.

[23] F. R. Gomez, A. Rajapakse, U. Annakkage, and I. Fernando, "Support vector machine-based algorithm for post-fault transient stability status prediction using synchronized measurements," *IEEE Trans. Power Syst.*, vol. 26, no. 3, pp. 1474–1483, Aug. 2011.

[24] H. Nguyen Duc, I. Kamwa, L.-A. Dessaint, and H. Cao-Duc, "A novel approach for early detection of impending voltage collapse events based on the support vector machine," *Int. Trans. Electr. Energy Syst.*, vol. 27, no. 9, p. e2375, Jun. 2017.

[25] (May 2008). *Rattle (the R Analytical Tool to Learn Easily), by D. Williams, Version 5.4*. [Online]. Available: http://rattle.togaware.com/

[26] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. New York, NY, USA: Springer-Verlag, 2008.

[27] H. Choi, J. Gim, Y. Seo, and D. Baik, "VPL-based big data analysis system: UDAS," *IEEE Access*, vol. 6, pp. 40883–40897, 2018.

[28] B. H. Menze, B. M. Kelm, R. Masuch, U. Himmelreich, P. Bachert, W. Petrich, and F. A. Hamprecht, "A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data," *BMC Bioinf.*, vol. 10, no. 1, p. 213, 2009.

[29] C. Chen, T. R. Razak, and J. M. Garibaldi, "FuzzyR: An extended fuzzy logic toolbox for the r programming language," in *Proc. IEEE Int. Conf. Fuzzy Syst. (FUZZ-IEEE)*, Jul. 2020, pp. 1–8.

[30] *Long Short-Term Memory Networks*. Accessed: Apr. 15, 2021. [Online]. Available: https://www.mathworks.com/help/deeplearning/ug/long-short-term-memory-networks.html

[31] C. Canizares, T. Fernandes, E. Geraldi, L. Gerin-Lajoie, M. Gibbard, I. Hiskens, J. Kersulis, R. Kuiava, L. Lima, F. DeMarco, N. Martins, B. C. Pal, A. Piardi, R. Ramos, J. dos Santos, D. Silva, A. K. Singh, B. Tamimi, and D. Vowles, "Benchmark models for the analysis and control of small-signal oscillatory dynamics in power systems," *IEEE Trans. Power Syst.*, vol. 32, no. 1, pp. 715–722, Jan. 2017.

[32] V. Kecman, *Support Vector Machines: Theory and Applications*. Berlin, Germany: Springer-Verlag, 2005.

[33] M. Pavella, D. Ernst, and D. Ruiz-Vega, *Transient Stability of Power Systems: A Unified Approach to Assessment and Control*. New York, NY, USA: Springer, 2000.

[34] S. M. Mazhari, B. Khorramdel, C. Y. Chung, I. Kamwa, and D. Novosel, "A simulation-based classification approach for online prediction of generator dynamic behavior under multiple large disturbances," *IEEE Trans. Power Syst.*, vol. 36, no. 2, pp. 1217–1228, Mar. 2021.

RAOULT TEUKAM DABOU (Student Member, IEEE) received the B.S. degree from the Faculty of Sciences, University of Ngaoundere, Cameroon, in 2009, and the M.S. degree in computer science, electronic, and automation from the University of Ngaoundere, in 2012. He is currently pursuing the Ph.D. degree in electrical and computer engineering with Laval University, Quebec City. His research interests include power dynamique stability, including intelligent control in control and power systems.

INNOCENT KAMWA (Fellow, IEEE) received the Ph.D. degree in electrical engineering from Université Laval, in 1989. He was a Researcher at the Hydro-Québec's Research Institute, specializing in the dynamic performance and control of power systems. He was the Chief Scientist of Hydro-Québec's Smart Grid Innovation Program and an international consultant in power grid simulation and network stability. He is currently a Full Professor with the Department of Electrical Engineering and the Tier 1 Canada Research Chair in decentralized sustainable electricity grids for smart communities at Laval University. He is a fellow of the Canadian Academy of Engineering. He was a recipient of the 2019 IEEE Charles Proteus Steinmetz and Charles Concordia Awards.

C. Y. CHUNG (Fellow, IEEE) received the B.Eng. (Hons.) and Ph.D. degrees in electrical engineering from The Hong Kong Polytechnic University, Hong Kong, in 1995 and 1999, respectively. He is currently a Professor, the NSERC/SaskPower (Senior) Industrial Research Chair at Smart Grid Technologies, and the SaskPower Chair in power systems engineering with the Department of Electrical and Computer Engineering, University of Saskatchewan, Saskatoon, SK, Canada.

CHUMA FRANCIS MUGOMBOZI (Member, IEEE) received the applied M.Sc. and Ph.D. degrees in electrical engineering from the École Polytechnique, Montréal, QC, Canada, in 2007 and 2013, respectively. He has been working as a Researcher with the Research and Innovation—Transmission Division, Systems Simulation and Evolution Group, Hydro-Québec Research Institute (IREQ), Varennes, QC, Canada, since 2010. He has been involved in modeling and simulation of HVDC and power systems simulator development, including parallel computation, nonlinear control equations solution, and co-simulation.