A Super-resolution Perception-based Incremental Learning Approach for Power System Voltage Stability Assessment with Incomplete PMU Measurements

Chao Ren, *Student Member, IEEE*, Yan Xu, *Senior Member, IEEE*, Junhua Zhao, *Senior Member, IEEE*, Rui Zhang, and Tong Wan

***Abstract*—This paper develops a fully data-driven, missing-data tolerant method for post-fault short-term voltage stability (STVS) assessment of power systems against the incomplete PMU measurements. The super-resolution perception (SRP), based on a deep residual learning convolutional neural network, is employed to cope with the missing PMU measurements. The incremental broad learning (BL) is used to rapidly update the model to maintain and enhance the online application performance. Being different from the state-of-the-art methods, the proposed method is fully data-driven and can fill up missing data under any PMU placement information loss and network topology change scenario. Simulation results demonstrate that the proposed method has the best performance in terms of STVS assessment accuracy and missing-data tolerance among the existing methods on the benchmark testing system.**

***Index Terms*—Data-driven, deep residual convolutional neural network, incremental broad learning, short-term voltage stability, super-resolution perception.**

I. Introduction

SHORT-TERM voltage stability (STVS) of a power system refers to the ability of the system to rapidly recover its bus voltages to an acceptable level following a large disturbance [1].

STVS is becoming a critical threat for power system security especially because of the increasing penetration of wind power resources. E.g., the 2016 South Australia blackout was caused by successive voltage disturbances which triggered a wind farm outage [2]. A data-driven scheme has been identified as a promising approach to achieve real-time post-fault STVS assessment based on phasor measurement unit (PMU) measurements. The principle is to train the machine learning (ML) model from a STVS database at the offline stage. For an application, a well-trained ML-based STVS model can work real-time with online PMU measurements, providing merits with less data requirements, much faster assessment speed, and better generalization capability [3].

The traditional model-based STVS assessment method is the time-domain simulation (TDS), which aims to solve the high-dimensional differential-algebraic equations with the complete and accurate information of system models, contingencies, and operating conditions. Due to the high computational burden of TDS, it is not feasible for post-fault STVS assessment, since short-term stability will be lost within a short period of time before the TDS can be completed. Thus, the TDS can neither make the decisions for real-time STVS assessment nor make the accurate decisions with the incomplete PMU data. To achieve the real-time assessment, several data-driven STVS methods are designed. In practice, the data-driven stability assessment can be divided into pre-fault assessment and post-fault assessment [4]. The former is to evaluate the power system stability status under a potential (but not yet occurred) disturbance, so the data inputs into the ML model are steady-state operational variables, such as power generation, load demand, bus voltage, and line flows, etc. The latter aims to timely predict the stability status after the fault occurrence. Therefore, the data inputs to the model are the post-fault dynamic trajectories, such as voltage and rotor angle trajectories.

In the literature, a variety of data-driven models have been reported with promising performance. In [5], a hierarchical method is proposed for pre-fault STVS assessment, which first classifies the voltage collapse status and then quantifies the voltage stability degree. In [6], a time series shapelet classification method is proposed for post-fault STVS assessment. A support vector machine (SVM) is used for transient stability [7] and voltage stability assessment [8]. A decision tree (DT) is used to predict the voltage stability state [9]. To address the imbalance problem in the training dataset, Refe-

Manuscript received November 8, 2020; revised February 8, 2021; accepted February 26, 2021. Date of online publication April 30, 2021; date of current version July 2, 2021. The work was supported in part by National Natural Science Foundation of China (51807009, 71931003, 72061147004).

C. Ren is with the Interdisciplinary Graduate School, Nanyang Technological University, Singapore.

Y. Xu is with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore.

J. H. Zhao is with the School of Science and Engineering, The Chinese University of Hong Kong (Shenzhen), China.

R. Zhang (corresponding author, email: rachelzhang.au@gmail.com; ORCID: https://orcid.org/0000-0003-1516-9653) is with Changsha University of Science and Technology, Changsha 410114, China.

T. Wan is with School of Electrical and Information Engineering, University of Sydney, Sydney, Australia.

DOI: 10.17775/CSEEJPES.2020.05930

rence [10] proposed an imbalance learning method. In [11], a time-adaptive scheme is proposed for faster post-fault STVS assessment without compromising the STVS accuracy. In [12], a hybrid ensemble learning approach is developed to further balance the accuracy and the speed for STVS assessment. In addition, recurrent neural networks (RNNs) are also utilized, since they have a better ability to consider spatial and temporal correlations, such as long short-term memory (LSTM) units [13], [14] and gated recurrent units (GRUs) [15].

All the above data-driven studies assumed that the model inputs are complete and available online. However, due to many practical issues, the data inputs may be missing, such as PMU failure, PDC failure, communication latency, or even cyber-attacks, etc. According to the wide-area measurements (WAMS) reliability analysis in [16], the WAMS reliability of a 14-bus system is assessed as 97.35%, that is, the occurrence probability of data missing is 2.65%. For the large-scale systems, the data missing probability may increase due to more extreme scenarios. With such missing real-time measurements, the existing methods will become ineffective due to the lack of model inputs.

In the literature, some methods have been proposed to alleviate the detrimental impact of missing data. For traditional stability assessment methods with incomplete input, References [17], [18] aim to fill up the missing data by using the available data, which suffers from heavy computational complexity and is improper for real-time post-fault assessment. In [19], a degenerate decision tree (DDT) uses a large number of single DTs, but its performance is limited by its tree structure. Through combining basic DTs trained via location-separated random subspace, DDT can achieve a higher accuracy performance compared with single DT methods. In [20], a decision tree with surrogate splits (DTSS) employs the surrogate splitting rule to substitute all the missing inputs, but suffers from imperfect accuracy performances because of the incomplete information. A random forest with surrogate splits (RFSS) [21] can be regarded as an ensemble model of DTSS in order to increase its diversity. All of the above studies focus on missing data imputation and always suffer from serious performance degradation once the number of missing data is increased. In [22], a mean imputation (MI) can substitute the missing features via the direct mean value of those features. In [23] and [24], robust feature ensemble learning (RFEL) is proposed to strategically collect observation-constrained PMU clusters as the training data, but a large number of well-trained classifiers are needed to use against all types of PMU missing scenarios, which would suffer from dimensional explosion.

However, it can be seen that most of the existing methods [19]–[23], and the generative adversarial network-based method [25] are for pre-fault stability assessment; only [24] considers the STVS with missing data conditions, and it cannot maintain a satisfactory performance if the amount of missing data increases. In addition, most of missing data recovering methods are fully dependent on the network topologies and PMU observability; once they are changed or the location of the missing PMU measurements are unknown, the model may become ineffective. All of the above data-driven stability methods do not consider real-time updating for trained stability models; once they need to use real-time updating or the IS cannot achieve the practical standards in advance, the existing trained data-driven stability models may become ineffective and have to be updated via re-training the whole model with the new parameter settings or a larger database, which will consume much more computation time.

In this paper, aiming to maintain the satisfactory performance with missing PMU measurements for post-fault STVS assessment, a novel data-driven missing-data tolerant method is proposed. The proposed method is composed of two models, including a missing data recovering model and a STVS prediction model with faster updating capability. The main contributions and values of this paper are as follows:

1) The missing data recovering model is developed based on the super-resolution perception (SRP) algorithm. Such a SRP model works for all types of data missing conditions with the same number of missed PMUs and is independent on PMU observability and network topologies. Also, such a SRP model can still accurately recover the missing data when a large amount of data is lost.

2) The STVS prediction model can be updated online for a real-time STVS application based on broad learning (BL) with an increment of feature nodes. The main value of the BL based model is its flexible structure, which allows much faster (online) model updating for performance enhancement and maintenance without retraining the whole model. Therefore, the proposed data-driven method is more generalized and extensible than the existing methods.

The proposed data-driven method is tested on the New England 10-machine 39-bus system, showing a higher STVS accuracy performance than existing methods. Through providing greater flexibility for system operators to manipulate the STVS assessment performance, the proposed method can be employed to substitute different incomplete PMU measurement scenarios.

II. Proposed Methodology

The implementation of the proposed IS requires two stages, an offline training process and an online application process. The proposed novel data-driven model based on SRP and incremental BL is displayed as Fig. 1.

A. Offline Training Process

In the offline training process, a historical STVS database is utilized to train a SRP model for recovering missing PMU data and a STVS model with faster updating capability based on incremental BL.

The SRP technique takes the incomplete data as the input and directly outputs the estimated complete data via feature extraction, feature supplement, and reconstruction. Therefore, the SRP model can accurately recover real-time complete PMU measurements from missing data (e.g., caused by PMU failures, PDC failure, communication latency, etc.) without adding any new data while ensuring the data quality is as high as possible. For the SRP model, the input is the incomplete voltage trajectories and the output is the recovered complete voltage trajectories.

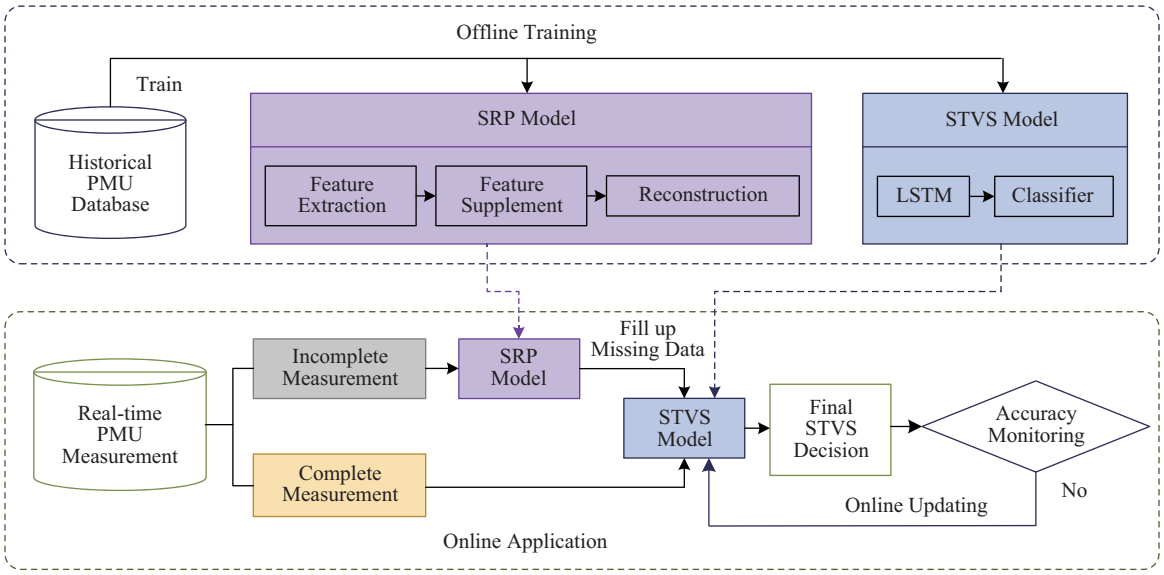

Fig. 1. The framework of the proposed method.

The STVS model can be a predictor/classifier trained via any effective ML algorithm. The post-fault STVS assessment can be regarded as a time series classification issue. In order to achieve online updating, the STVS model is trained as a classifier based on the incremental BL algorithm, which is suitable for time series classification. For the STVS model, the input is complete voltage trajectory measurements and the output is the stability status (i.e., stable or unstable).

B. Online Application Process

For the online application process, the real-time PMU measurements can be used as model inputs. The proposed method first estimates whether the real-time measurements are complete. If the real-time measurements are complete, they will be directly sent to the STVS model; otherwise, the incomplete data will be imported into the SRP model to recover the missing PMU measurements. Then, the complete PMU data can be sent to the trained STVS model. With such complete measurement data, the STVS model is able to make an accurate and timely decision on the system's stability status. If the accuracy monitoring judges that it does not meet the actual STVS assessment accuracy requirements, the STVS model needs to update online, then the current STVS model can be continuously updated by the increments of the features for performance maintenance or improvement. Finally, when the system is estimated to be unstable, then emergency control actions should be activated, such as load shedding.

III. Methodologies

The proposed data-driven method is based on SRP and BL with increments of feature nodes. This section first introduces the principle of SRP, then describes the fundamentals of BL and its incremental scenario.

A. Super-Resolution Perception (SRP)

The SRP technique is designed to generate high-resolution data D_{complete} from low-resolution missing data D_{missing} [26]. Both D_{complete} and D_{missing} represent the value of the same physical quantity in the common time period, and are related by a degradation model as shown in (1):

$$D_{\text{missing}} =\downarrow D_{\text{complete}} + noise \tag{1}$$

where $\downarrow$ represents the degradation function and noise represents the noise. SRP can be regarded as the inference D_{complete} with D_{missing} by the SRP mapping function $R_\theta(\cdot)$ as shown in (2):

$$D'_{\text{complete}} = R_\theta\left(D_{\text{missing}}\right) \tag{2}$$

where the mapping function $R_\theta(\cdot)$ can recover the information missed by degradation function $\downarrow$ as much as possible. In order to to find the best possible solution of SRP, an additional constraint called Maximum a Posteriori estimation (MAP) is considered, where the final generated D'_{complete} is the solution with maximum posterior probability $p(D_{\text{complete}} \mid D_{\text{missing}})$. Based on the Bayesian formula, the corresponding D_{complete} given the specific D_{missing} can be calculated via solving the MAP estimation as shown in (3):

$$\begin{aligned} D'_{\text{complete}} &= \arg\max_{D_{\text{complete}}} p\left(D_{\text{complete}} \mid D_{\text{missing}}\right) \\ &= \arg\max_{D_{\text{complete}}} p\left(D_{\text{missing}} \mid D_{\text{complete}}\right) p\left(D_{\text{complete}}\right) \end{aligned} \tag{3}$$

where $p\left(D_{\text{missing}} \mid D_{\text{complete}}\right)$ is the likelihood and can be obtained by the degradation model; $p\left(D_{\text{complete}}\right)$ can be solved by the prior model as the regularization term to constrain the solution of the estimation D'_{complete} satisfying the prior distribution of D_{complete}. Equation (3) is equal to (4) as follows:

$$\begin{aligned} D'_{\text{complete}} = \arg\max_{D_{\text{complete}}} &\log p\left(D_{\text{missing}} \mid D_{\text{complete}}\right) + \\ &\log p\left(D_{\text{complete}}\right) \end{aligned} \tag{4}$$

The modeling of the prior D_{complete} has an important impact on the SRP problem. Through effectively modeling the prior of high-resolution data D_{complete} and degradation process,

D_{complete} can be well estimated. Then, the high-resolution data D_{complete} is estimated through the deep neural networks given the low-resolution missing data D_{missing}. The purpose of the network is to minimize the mean squared error loss function as shown in (5):

$$\begin{aligned} L\left(D_{\text{complete}}, D'_{\text{complete}}\right) &= \left\|D_{\text{complete}} - D'_{\text{complete}}\right\|_2^2 \\ &= \left\|D_{\text{complete}} - R_\theta\left(D_{\text{missing}}\right)\right\|_2^2 \end{aligned} \quad (5)$$

Then, the optimal model parameters θ can be calculated via minimizing the loss function (5) with the gradient descent algorithm as shown in (6):

$$\theta' = \arg\min_\theta \left\|D_{\text{complete}} - R_\theta\left(D_{\text{missing}}\right)\right\|_2^2 \quad (6)$$

The SRP mapping function $R_\theta(\cdot)$ is implemented via the convolutional neural network. The network applies the incomplete data D_{missing} as the input features with the length of D_f and directly outputs the approximated complete data D'_{missing} with the length of D_c, $D_f < D_c$. The structure of SRP is shown in Fig. 2, and SRP is composed of three parts, including feature extraction, feature supplement, and reconstruction.

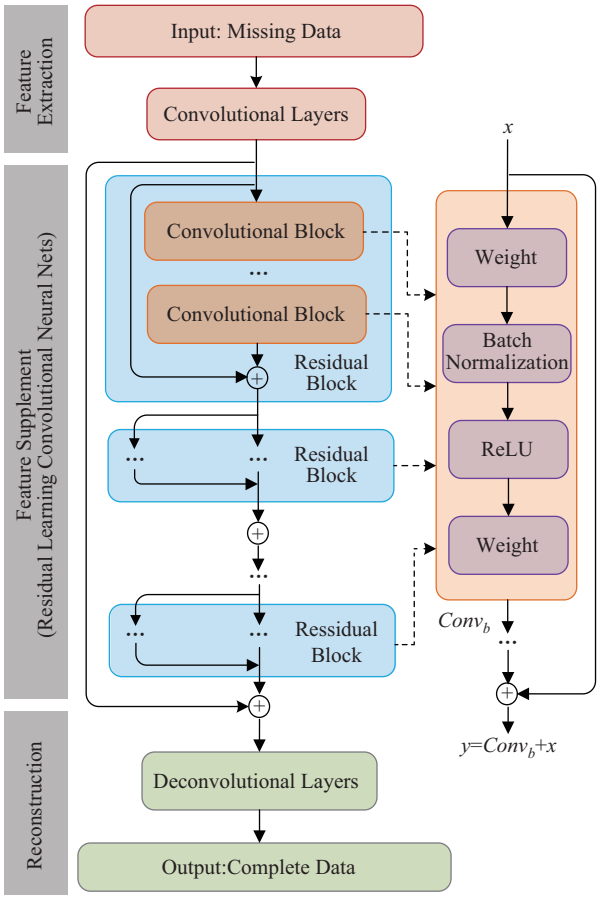

Fig. 2. The structure of SRP.

The beginning of the convolution layers is applied for a global feature extractor, which can extract abstract feature information from the D_{missing} and handle the features as many of the same dimension feature vectors. Given inputs $\boldsymbol{X} \in \mathbb{R}^{p \times d_f}$ with q instances and d_f features, the feature extraction portion extracts features from input $\boldsymbol{X}$ and all the samples are represented as the v feature vectors with the length of d_f. These features $\boldsymbol{F} \in \mathbb{R}^{q \times v \times d} f$ represent the abstract feature information of input $\boldsymbol{X}$. With such extracted features, an information supplemental sub-network based on a deep residual learning convolutional neural network (ResNet) [27] is utilized to supplement the missing information, which adds the local connection in each residual block to ensure the gradient flow directly through the bottom layers. Under such a structure, the global connection can force learning on the missed information. A deep ResNet is composed of a number of residual blocks. The structure of the residual block is composed of several convolutional blocks Convs with the number of filter b, including convolutional layer, batch normalization layer and the Rectified Linear Units (ReLU) activation function. Both the local residual connections and the global residual connections are applied to improve the accuracy performance via residual functions. After that, the reconstruction part can integrate the feature vectors of samples into α sub-vectors with the length of d_f. Then, the sub-vectors $\boldsymbol{F} \in \mathbb{R}^{q \times \alpha \times d_f}$ are rearranged as the estimated complete data $\boldsymbol{Y} \in \mathbb{R}^{q \times d_c}$. It should be noted that the sub-vectors can be generated in parallel via applying the convolution operation in parallel, since the proposed SRP method has a heavy computational efficiency.

As mentioned above, the output of the SRP model is the estimated complete data which is reconstructed from the original data. Thus, the estimated complete data can accurately represent the characteristics of the original data based on the effective feature learning ability of the *ResNet*.

B. Broad Learning (BL) Algorithm

The BL algorithm shown in Fig. 3 is derived from a random vector functional link network [28]. It can obtain high accuracy performance without deep neural network structures, while the deep multi-layer structures suffer from huge computational burden and need much more time to train a large number of parameters. Its structure consists of input $\boldsymbol{X}$, features $\boldsymbol{Z}^n$, enhancement hidden nodes $\boldsymbol{H}^m$ and output $\boldsymbol{Y}$. First, the input $\boldsymbol{X}$ is transformed into feature nodes $\boldsymbol{Z}^n$ via feature mappings, and then the feature mapping is connected via the nonlinear activation function to form the enhancement hidden nodes $\boldsymbol{H}^m$. Finally, features $\boldsymbol{Z}^n$ and the corresponding enhancement hidden nodes $\boldsymbol{H}^m$ are connected to the output $\boldsymbol{Y}$, and then the output weights $\boldsymbol{W}$ can be calculated via the regression way.

$$\boldsymbol{Y} = \boldsymbol{R}_n^m \boldsymbol{W}_n^m = \left[\boldsymbol{Z}^n(\boldsymbol{X}) \mid \boldsymbol{H}^m(\boldsymbol{X})\right] \boldsymbol{W}_n^m \quad (7)$$

where

$$\begin{aligned} \boldsymbol{R}_n^m &= \left[\boldsymbol{Z}^n(\boldsymbol{X}) \mid \boldsymbol{H}^m(\boldsymbol{X})\right] \\ \boldsymbol{Z}_i &= \left[\mu\left(\boldsymbol{X}\boldsymbol{W}_{e_i} + \boldsymbol{\beta}_{e_i}\right)\right], i = 1, 2, \ldots, n \\ \boldsymbol{Z}^n &= \left[\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n\right] \\ \boldsymbol{H}_j &= \left[\varphi\left(\boldsymbol{Z}^n \boldsymbol{W}_{h_j} + \boldsymbol{\beta}_{h_j}\right)\right], j = 1, 2, \ldots, m \\ \boldsymbol{H}^m &= \left[\boldsymbol{H}_1, \ldots, \boldsymbol{H}_m\right] \end{aligned} \quad (8)$$

where state matrix is $\boldsymbol{R}_n^m = \left[\boldsymbol{Z}^n(\boldsymbol{X}) \mid \boldsymbol{H}^m(\boldsymbol{X})\right]$; n and m denote the quantity of features and enhancement hidden nodes.

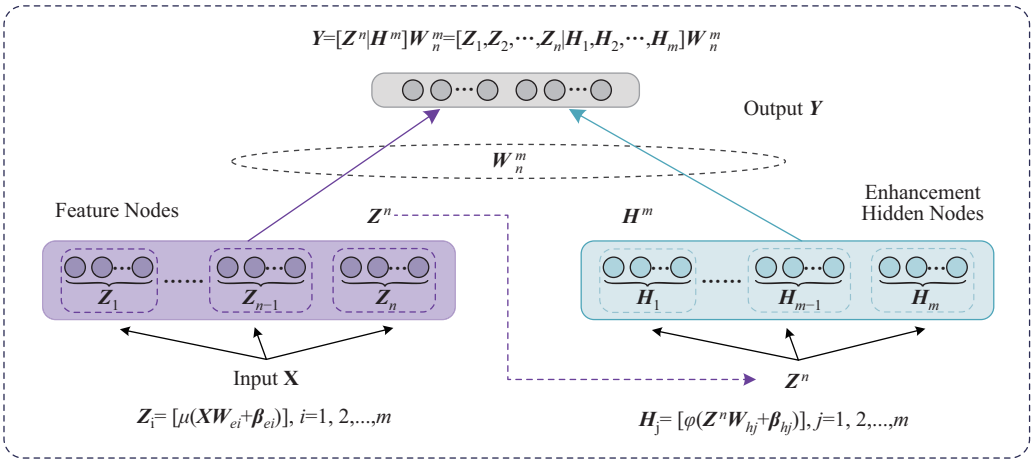

Fig. 3. The structure of the basic BL algorithm.

C. Increment of Feature Nodes

For the practical real-time application, the initially chosen significant features might not be sufficient to represent the real-time power systems stability status in a timely manner. In order to solve such a problem, the traditional method is to utilize more features to retrain the model and regain the model parameters for the larger and deeper neural networks, which would need more time and more computational resources.

For BL, the whole model is efficiently constructed and does not need to retrain the whole network. The BL networks can be expanded via adding additional features to obtain the latest state matrix $\boldsymbol{R}_{n+1}^m$ as follows:

$$\boldsymbol{R}_{n+1}^m = [\boldsymbol{R}_n^m \,|\boldsymbol{Z}_{n+1}|\, \boldsymbol{H}_{ex_m}] \tag{9}$$

where $\boldsymbol{H}_{ex_m}$ represents additional output of the enhancement hidden node which corresponds to the $(n+1)$-th group of features $\boldsymbol{Z}_{n+1}$. With the increment of features and the enhancement hidden nodes, the latest output weights $\boldsymbol{W}_{n+1}^m$ can be updated by Eq. (10):

$$\boldsymbol{W}_{n+1}^m = \left(\boldsymbol{R}_{n+1}^m\right)^+ \boldsymbol{Y} = \begin{bmatrix} \boldsymbol{W}_n^m - \boldsymbol{V}\boldsymbol{K}^T\boldsymbol{Y} \\ \boldsymbol{K}^T\boldsymbol{Y} \end{bmatrix} \tag{10}$$

where the pseudoinverse of $\boldsymbol{R}_{n+1}^m$ is as shown in (11).

$$\left(\boldsymbol{R}_{n+1}^m\right)^+ = \begin{bmatrix} \boldsymbol{R}_n^m - \boldsymbol{V}\boldsymbol{K}^T \\ \boldsymbol{K}^T \end{bmatrix} \tag{11}$$

where

$$\begin{aligned} \boldsymbol{V} &= (\boldsymbol{R}_n^m)^+ [\boldsymbol{Z}_{n+1} \mid \boldsymbol{H}_{ex_m}] \\ \boldsymbol{P} &= [\boldsymbol{Z}_{n+1} \mid \boldsymbol{H}_{ex_m}] - \boldsymbol{R}_n^m \boldsymbol{V} \\ \boldsymbol{K}^T &= \begin{cases} \boldsymbol{P}^+ & \boldsymbol{P} \neq \boldsymbol{0} \\ \left(\boldsymbol{1} + \boldsymbol{V}^T\boldsymbol{V}\right)^{-1} \boldsymbol{V}^T (\boldsymbol{R}_n^m)^+ & \boldsymbol{P} = \boldsymbol{0} \end{cases} \end{aligned} \tag{12}$$

It can be seen that the BL with the increment of features only needs to compute the output weights related to the $(n+1)$-th group of features, thus leading to a highly computational burden. The structure of BL with the increment of features scenario is shown in Fig. 4, and Algorithm 1 provides the complete pseudo-code of this learning procedure.

Algorithm 1: Incremental BL

Input : Input samples $\boldsymbol{X}$, the quantity of features n, the quantity of enhancement hidden nodes group m.
Output: Output Weight $\boldsymbol{W}$.
begin
  for $i = 1$ *to* n **do**
    Randomly set $\boldsymbol{W}_{e_i}$, $\boldsymbol{\beta}_{e_i}$.
    Calculate $\boldsymbol{Z}_i = [\mu(\boldsymbol{X}\boldsymbol{W}_{e_i} + \boldsymbol{\beta}_{e_i})]$.
  end
  Get the features group $\boldsymbol{Z}^n = [\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n]$.
  for $j = 1$ *to* m **do**
    Randomly set $\boldsymbol{W}_{h_j}$, $\boldsymbol{\beta}_{h_j}$.
    Calculate $\boldsymbol{H}_j = [\varphi(\boldsymbol{Z}^n\boldsymbol{W}_{h_j} + \boldsymbol{\beta}_{h_j})]$.
  end
  Get the enhancement hidden nodes group $\boldsymbol{H}^m = [\boldsymbol{H}_1, \ldots, \boldsymbol{H}_m]$.
  Set $\boldsymbol{R}_n^m$ and calculate $(\boldsymbol{R}_n^m)^+$ by Eq. (7), (8).
  repeat
    # *Increment of features*
    Randomly set $\boldsymbol{W}_{e_{n+1}}$, $\boldsymbol{\beta}_{e_{n+1}}$.
    Calculate $\boldsymbol{Z}_{n+1} = [\mu(\boldsymbol{X}\boldsymbol{W}_{e_{n+1}} + \boldsymbol{\beta}_{e_{n+1}})]$.
    Randomly set $\boldsymbol{W}_{ex_i}$, $\boldsymbol{\beta}_{ex_i}$, i=1, …, m.
    Calculate $\boldsymbol{H}_{ex_m} =$
    $[\varphi(\boldsymbol{Z}_{n+1}\boldsymbol{W}_{ex_i}+\boldsymbol{\beta}_{ex_i}), \ldots, \varphi(\boldsymbol{Z}_{n+1}\boldsymbol{W}_{ex_m}+\boldsymbol{\beta}_{ex_m})]$.
    Update $\boldsymbol{R}_{n+1}^m$ by Eq. (7).
    Compute $(\boldsymbol{R}_{n+1}^m)^+$ and $\boldsymbol{W}_{n+1}^m$ by Eq. (9), (10), (11), (12).
    $n = n + 1$
  until *Satisfy the error threshold*
end

IV. Simulation Results

The proposed approach is tested on a modified New England 10-machine 39-bus system as shown in Fig. 5 [23], and the corresponding PMU placement locations are listed in Table I. With the increasing integration of active demand-side

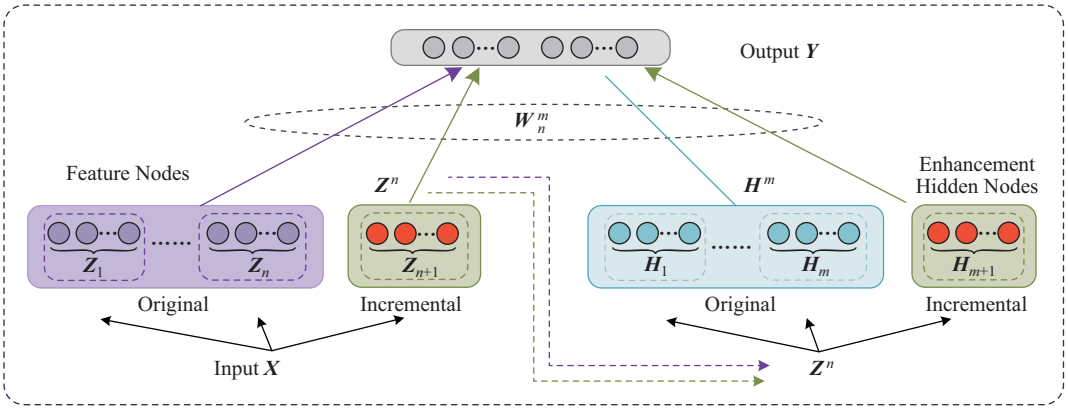

Fig. 4. The structure of the BL with increment of features.

TABLE I
PMU PLACEMENT LOCATION

Placement	PMU installation buses	Number of PMUs
■	3, 8, 10, 16, 20, 23, 25, 29	8

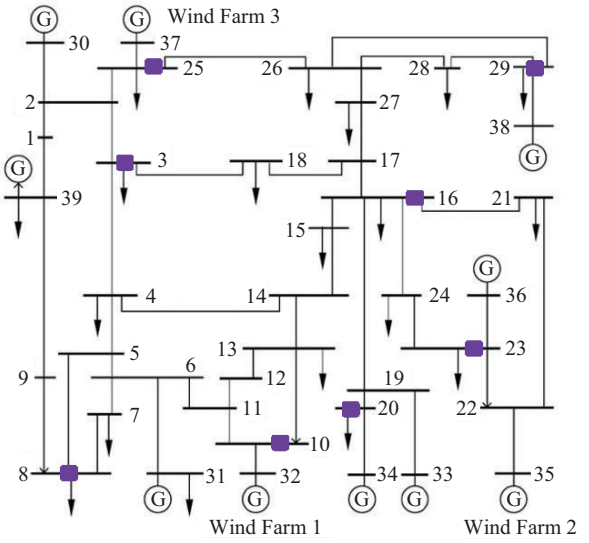

Fig. 5. The modified New England 39-bus system.

response and renewable energy sources, more uncertainties are introduced into the power systems. Considering the effect of renewable energy resources on power systems, three wind farms replace the synchronous generators on buses 32, 35, and 37. The capacities of other generators are increased by 1.5 times to accommodate the wind power, and the wind power penetration of the modified system is 36.58%. The simulation is conducted on a computer with an Intel Core i7 CPU @ 3.3-GHz processor, 16-GB RAM and GPU with Nvidia GeForce GTX 1060. Time-domain simulations are performed via the commercial software PSS/E [29]. The proposed method is implemented based on Python 3.6 with the Pytorch framework.

A. Database Generation Process

A comprehensive STVS database is generated to consider a wide range of system operating scenarios and post-disturbance voltage trajectories. In this paper, the uncertainties of factors need to be carefully considered, including pre-fault operating points, dynamic load models and fault scenarios, etc., which are respectively determined as follows.

- *Pre-fault Operating Points:* Through randomly varying the load demand between 0.8 and 1.2 times that of the base values, and the wind power output between 0 and its capacity, 6,536 operating points are generated. In addition, optimal power flow is run to determine the synchronous generators' output.
- *Fault Scenarios:* The three-phase short-circuit faults are considered in the simulation with a random fault duration between 0.1 s and 0.3 s. Then, a randomly chosen fault location is employed for each operating point. According to the practical scenarios, the fault should be cleared either with a single transmission line tripping or without loss of the power grid component to simulate the different fault-induced topology changes.
- *Load Modeling:* "CLOD" [29] applied in PSS/E is utilized for the testing, which is an industry-standard composite load model. The "CLOD" model is composed of 6 typical load components in practical substations, including transformer saturation, discharge lighting, small motors, large motors, and voltage-dependent loads. The portion of motor loads for all the operating points is randomly sampled between 0% and 80%. Through measurement-based load modeling methods [30], the different load components share can be obtained.

The distribution of the transient voltage severity index (TVSI) value and load level are displayed in Fig. 6. The load level represents the range of the operating systems, and the TVSI proposed in [5] is utilized to evaluate the severity of the post-fault voltage deviation. The small TVSI value means fast voltage recovery, while a large TVSI value represents unstable voltage propagation. For the generated database, the quantity of the stable samples and unstable samples are 4,023 and 2,513, respectively. 80% of the samples are randomly chosen for the training process, and the remaining 20% samples for testing.

B. Observation Windows

In general, the longer the observation windows, the more measurements can be obtained, so the STVS assessment results

TABLE II
SNAPSHOT PERFORMANCES OF STVS ACCURACY AND COMPUTATIONAL EFFICIENCY BY INCREMENT OF FEATURES FOR STVS DATABASE

No. of training samples	Time period length as feature nodes	No. of enhancement nodes	Testing accuracy	Each additional training times	Accumulative training times	Each additional testing times	Accumulative testing times
5228 (base case)	0.8 s	100	96.19%	0.1387 s	0.1387 s	0.0215 s	0.0215 s
5228	0.8 s ⟶ 0.9 s	100 ⟶ 150	96.98%	0.0571 s	0.1958 s	0.0058 s	0.0273 s
5228	0.9 s ⟶ 1.0 s	150 ⟶ 200	97.53%	0.0576 s	0.2534 s	0.0051 s	0.0324 s
5228	1.0 s ⟶ 1.1 s	200 ⟶ 250	98.09%	0.0643 s	0.3177 s	0.0052 s	0.0376 s
5228	1.1 s ⟶ 1.2 s	250 ⟶ 300	98.37%	0.0678 s	0.3855 s	0.0049 s	0.0425 s

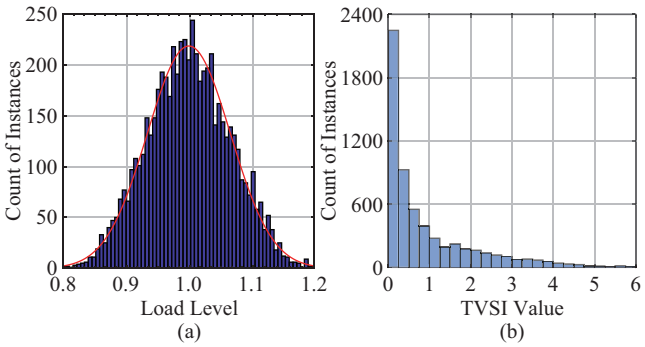

Fig. 6. The distribution of the load level and TVSI value for generated STVS database. (a) load level. (b) TVSI.

tend to be more accurate. However, if the observation windows are too long, the system will lose the stability and cannot activate emergency control timely, and instability cannot be avoided [31]. Thus, it is necessary to predict the voltage stability status as early as possible in order to leave enough time for activating the emergency controls. As the New England 10-machine 39-bus system has a nominal frequency of 60 Hz, the time interval of PMU data reporting is set to 1/60 s, 1/30 s, etc. To be more practical, the numerical tests have been re-performed with a 1/60 s assessment time interval.

Since this paper primarily aims at the tolerance of missing data, only fixed observation windows need to be considered. For the testing, three different and reasonable observation windows are chosen for STVS assessment, which are 0.8 s (48 sampling points), 1.0 s (60 sampling points) and 1.2 s (72 sampling points) after the fault clearance.

C. Increment of BL Testing Results

The snapshot results of the increment with feature nodes are shown in Table II. The columns of Table II, from left to right, list the data-driven STVS method, the quantity of training samples, the length of time period as features (each time point includes 39 bus voltage values), the quantity of enhancement hidden nodes, testing STVS assessment accuracy, each additional training time, the accumulative training time, each additional testing time, and the accumulative testing time respectively. As an example, the accumulative training time and each additional training time denote the whole training time and the training time at the current step, respectively; the accumulative testing time and each additional testing time denote the whole testing time and the testing time at the current increment step, respectively.

As the basic case, 5228 samples with 48 time points (0.8 s) are employed to train the STVS model, and the quantity of enhancement hidden nodes is set as 100, resulting in the 96.19% testing STVS assessment accuracy. With the incremental BL, the STVS model can effectively achieve the high STVS assessment accuracy after four times incremental steps. With the increment of feature scenarios, the testing accuracy can continuously increase, up to 98.37%. Moreover, it can be seen that for each incremental step, the computation time is very short, which can be further verify that the proposed method is very suitable for STVS online updating without retraining a larger data-driven STVS model.

TABLE III
AVERAGE STVS PERFORMANCE OF DIFFERENT DATA-DRIVEN METHODS UNDER THREE DIFFERENT OBSERVATION WINDOWS

Data-driven STVS methods	Average STVS performance under three observation windows (0.8 s, 1.0 s, 1.2 s)	
	Average accuracy	Average computation time
LSTM	98.41%	5.327 s
FCNN	97.83%	12.532 s
BPNN	97.21%	0.843 s
Proposed BL	98.38%	0.459 s

In order to demonstrate the high efficiency of BL, the proposed BL-based STVS model is compared with state-of-the-art data-driven models under three observation windows as shown in Table III, including LSTM, fully convolutional neural network (FCNN), and back-propagation neural network (BPNN). Note that all of the data-driven STVS models under comparison have been well trained and tuned for their best performances. As shown in Table III, it can be seen that the proposed BL-based STVS model can achieve the 98.38% STVS accuracy with the fastest computation time, only 0.459 s, which can achieve $\sim$ 10X speedup compared with the LSTM-based STVS model. Thus, the proposed method can achieve online updating without sacrificing STVS accuracy.

D. SRP Testing Performance

The proposed SRP method shows the better performance in terms of the following five indicators, compared with four existing data-driven methods to cope with missing PMU measurements, including MI [22], DTSS [20], RFSS [21], and RFEL [24]. Note that all the testing results are the average performance values for each of the missing PMU placement data conditions.

1) Performance of Filled Missing Data

First, the performance of the filled missing data by the SRP model is tested. The error is evaluated via dynamic time warping (DTW) [32] and mean average percentage error

TABLE IV
SRP PERFORMANCES

Error index	Average computation efficiency	Filled missing data performance (under different missing PMU numbers)						
		1	2	3	4	5	6	7
MAPE	2.79 ms	2.58%	2.73%	2.93%	3.36%	3.77%	3.92%	4.12%
DTW		0.0158	0.0193	0.0246	0.0291	0.0375	0.0459	0.0565

(MAPE) [33]. DTW is one of the time series analysis algorithms used to measure the similarity between two different time series. A smaller DTW value means a higher similarity of STVS between generated PMU data with real PMU data.

Table IV lists DTW and MAPE of SRP for all types of missing PMU conditions. It is clear that the estimated missing data is very close to the ground truth value, which validates that the generated estimated data via the SRP model can accurately recover the incomplete PMU measurements. In addition, the average computation time is 2.79 ms of one instance, which can meet the requirement of IEEE Standard C37.118.2-2011 [34] and is negligible for real-time STVS applications.

2) STVS Assessment Accuracy Comparison

The testing results corresponding to three different observation windows are demonstrated in Fig. 7(a–c), where the STVS assessment accuracy represents the percentage of the correctly classified testing instances, and each plotted point denotes the accuracy under each identical number of random missing PMU placement measurements.

According to Fig. 7(a–c), with the increasing of the percentage of missing data, the existing methods show the different degrees of decrease in STVS assessment accuracy. But the proposed method has not been significantly affected, and can still maintain a relatively satisfactory STVS assessment accuracy, verifying its anti-missing ability.

3) Average Accuracy with Incomplete PMU Measurement Comparisons

A new technical index, called average STVS assessment accuracy (ASAA) shown in (13), is proposed to evaluate the STVS assessment accuracy and robustness under all types of PMU missing conditions:

$$\text{ASAA} = \frac{1}{W} \cdot \frac{1}{M} \cdot \sum_{w=1}^{\text{W}} \sum_{\text{m}=1}^{\text{M}} A_m \times 100\% \tag{13}$$

where W is the quantity of different observation windows. A_m represents the STVS assessment accuracy with random m PMU missing. ASAA represents the average STVS assessment accuracy for incomplete PMU measurement situations. A larger ASAA value indicates higher STVS assessment accuracy and robustness against missing PMU measurements.

4) Slope of Accuracy Drop Comparison

The missing-data tolerance can be indicated by the average slope of accuracy drop (ASAD) as shown in (14), where a lower slope represents heavier missing-data tolerance.

$$\text{ASAD} = \frac{1}{W} \cdot \frac{1}{M} \cdot \sum_{w=1}^{\text{W}} \sum_{m=1}^{M} \frac{A_0 - A_m}{d_m} \times 100\% \tag{14}$$

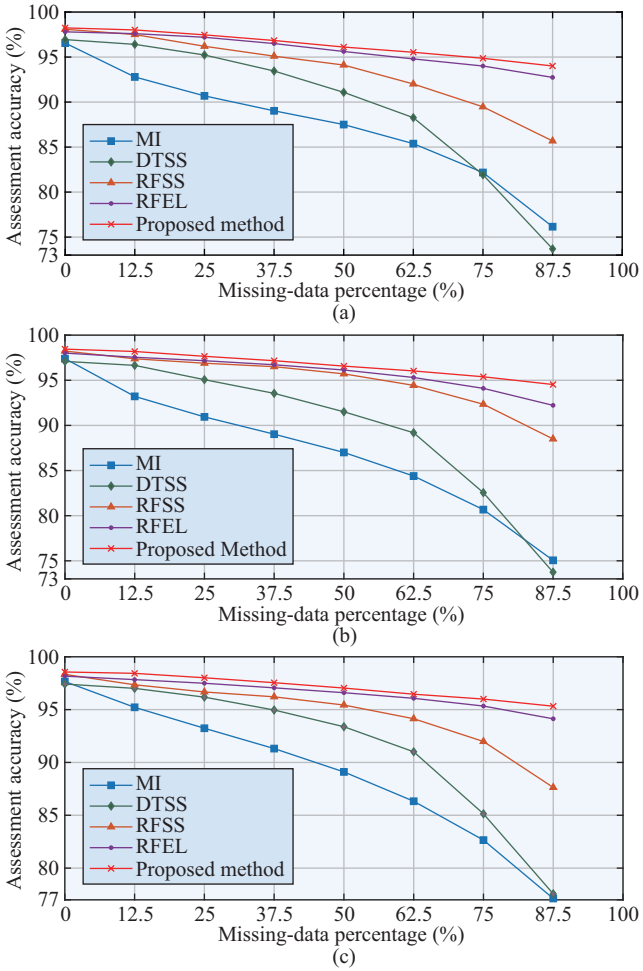

Fig. 7. STVS testing results among five different methods under three observation windows. (a) 0.8 s. (b) 1.0 s. (c) 1.2 s.

where A_0 represents the STVS assessment accuracy with complete PMU measurements, and d_m denotes the m-th missing data percentage value.

Five different methods of ASAA and ASAD shown in Fig. 7 are compared in Table V. It is clear that the proposed SRP method shows the best performance in terms of STVS assessment accuracy and missing-data tolerance among the

TABLE V
PERFORMANCE OF DIFFERENT METHODS AGAINST MISSING DATA FOR STVS DATABASE

Method	CEC	ASAA	ASAD
MI [22]	255	86.62%	0.2172
DTSS [20]	1	89.41%	0.1218
RFSS [21]	1	93.86%	0.0781
RFEL [24]	19	95.81%	0.0386
Proposed method	7	96.76%	0.0337

five different methods, therefore, the proposed method is demonstrated as having robustness against missing data.

5) Average Computational Efficiency Comparison

Less computational efficiency is also a significant merit of the proposed method, which is the number of classifiers needed for the whole operating system (CEC). The proposed method can recover the flawed PMU data under any missing condition through using less trained classifiers. However, in order to obtain the satisfactory results, some of existing methods need to train numerous classifiers (e.g., for M different PMUs, the total quantity of classifiers is $2^{\mathrm{M}} - 1$ at most). In Table V, it is obvious that the proposed data-driven method is more efficient than some of the methods which need to train more classifiers.

V. Conclusions, Discussions, and Future Works

In this paper, a novel data-driven missing-data tolerant approach based on SRP is proposed which can maintain the STVS assessment accuracy under any PMU missing conditions. In addition, in order to ensure online STVS applications, a novel BL algorithm with increments of feature nodes is applied, which allows it to rapidly expand broadly without retraining the whole ML-based STVS model due to its flexible structure. Compared with existing methods, the proposed method can accurately recover the missing PMU measurements without concern for the detailed missing PMU measurements placement to choose the model. Simulation results have validated its higher STVS assessment accuracy among the existing methods. Moreover, SRP can be significantly enhanced toward similar practical safety-critical issues in power engineering.

There are two main limitations of the proposed method. For the SRP model, there are many model parameters that need to be fine-tuned during the offline training stage, such as the network structure and the choice of the gradient descent algorithms, since it is based on the deep neural network with CNN. For the STVS model, this paper only considers the broad learning with the increments of features.

For future studies, first, a simpler model of network structure can be used, aiming to simplify the parameter adjustment process; secondly, in order to further improve the STVS assessment performance, the increment of instances scenario by incremental learning can be considered.

References

[1] P. Kundur, J. Paserba, V. Ajjarapu, G. Andersson, A. Bose, C. Canizares, N. Hatziargyriou, D. Hill, A. Stankovic, C. Taylor, T. Van Cutsem, and V. Vittal, "Definition and classification of power system stability IEEE/CIGRE joint task force on stability terms and definitions," *IEEE Transactions on Power Systems*, vol. 19, no. 3, pp. 1387–1401, Aug. 2004.

[2] Australian Energy Market Operator (AEMO). (2017, Mar.). Black system South Australia 28 Sep 2016, Final Report. [Online]. Available: https://www.aemo.com.au/.

[3] Z. Y. Dong, Y, Xu, P. Zhang, and K. P. Wong, "Using IS to assess an electric power system real-time stability," *IEEE Intelligent Systems*, vol. 29, no. 5, pp. 61–67, 2014.

[4] Y. Xu, Y. C. Zhang, Z. Y. Dong, and R. Zhang, *Intelligent systems for Stability Assessment and Control of Smart Power Grids*, Boca Raton: CRC Press, 2020.

[5] Y. Xu, R. Zhang, J. H. Zhao, Z. Y. Dong, D. H. Wang, H. M. Yang, and K. P. Wong, "Assessing short-term voltage stability of electric power systems by a hierarchical intelligent system," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 8, pp. 1686–1696, Aug. 2016.

[6] L. P. Zhu, C. Lu, and Y. Z. Sun, "Time series shapelet classification based online short-term voltage stability assessment," *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 1430–1439, Mar. 2016.

[7] B. Wang, B. W. Fang, Y. J. Wang, H. S. Liu, and Y. L. Liu, "Power system transient stability assessment based on big data and the core vector machine," *IEEE Transactions on Smart Grid*, vol. 7, no. 5, pp. 2561–2570, Sep. 2016.

[8] H. Mohammadi, G. Khademi, M. Dehghani, and D. Simon, "Voltage stability assessment using multi-objective biogeography-based subset selection," *International Journal of Electrical Power & Energy Systems*, vol. 103, pp. 525–536, Dec. 2018.

[9] H. Mohammadi and M. Dehghani, "PMU based voltage security assessment of power systems exploiting principal component analysis and decision trees," *International Journal of Electrical Power & Energy Systems*, vol. 64, pp. 655–663, Jan. 2015.

[10] L. P. Zhu, C. Lu, Z. Y. Dong, and C. Hong, "Imbalance learning machine-based power system short-term voltage stability assessment," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 5, pp. 2533–2543, Oct. 2017.

[11] Y. C. Zhang, Y. Xu, Z. Y. Dong, and R. Zhang, "A hierarchical self-adaptive data-analytics method for real-time power system short-term voltage stability assessment," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 1, pp. 74–84, Jan. 2019.

[12] C. Ren, Y. Xu, Y. C. Zhang, and R. Zhang, "A hybrid randomized learning system for temporal-adaptive voltage stability assessment of power systems," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 6, pp. 3672–3684, Jun. 2020.

[13] J. J. Q. Yu, D. J. Hill, A. Y. S. Lam, J. T. Gu, and V. O. K. Li, "Intelligent time-adaptive transient stability assessment system," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 1049–1058, Jan. 2018.

[14] L. Zheng, W. Hu, K. Y. Hou, X. W. Xu, and G. H. Shao, "Real-time transient stability assessment based on deep recurrent neural network," in *2017 IEEE Innovative Smart Grid Technologies-Asia: ISGT-Asia 2017*, Auckland, New Zealand, 2017, pp. 1–5.

[15] A. Gupta, G. Gurrala, and P. S. Sastry, "Instability prediction in power systems using recurrent neural networks," in *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI)*, 2017, pp. 1795–1801.

[16] Y. Wang, W. Y. Li, and J. P. Lu, "Reliability analysis of wide-area measurement system," *IEEE Transactions on Power Delivery*, vol. 25, no. 3, pp. 1483–1491, Jul. 2010.

[17] P. Z. Gao, M. Wang, S. G. Ghiocel, J. H. Chow, B. Fardanesh, and G. Stefopoulos, "Missing data recovery by exploiting low-dimensionality in power system synchrophasor measurements," *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 1006–1013, Mar. 2016.

[18] J. J. Q. Yu, A. Y. S. Lam, D. J. Hill, Y. H. Hou, and V. O. K. Li, "Delay aware power system synchrophasor recovery and prediction framework," *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 3732–3742, Jul. 2019.

[19] M. He, V. Vittal, and J. S. Zhang, "Online dynamic security assessment with missing pmu measurements: A data mining approach," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1969–1977, May 2013.

[20] T. Y. Guo and J. V. Milaović, "The effect of quality and availability of measurement signals on accuracy of on-line prediction of transient stability using decision tree method," in *Proceedings of 4th IEEE/PES ISGT Europe*, Lyngby, Denmark, 2013, pp. 1–5.

[21] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct. 2001.

[22] G. E. A. P. A. Batista and M. C. Monard, "An analysis of four missing data treatment methods for supervised learning," *Applied Artificial Intelligence*, vol. 17, no. 5–6, pp. 519–533, Nov. 2010.

[23] Y. C. Zhang, Y. Xu, R. Zhang, and Z. Y. Dong, "A missing-data tolerant method for data-driven short-term voltage stability assessment of power systems," *IEEE Transactions on Smart Grid*, vol. 10, no. 5, pp. 5663–5674, Sep. 2019.

[24] Y. C. Zhang, Y. Xu, and Z. Y. Dong, "Robust ensemble data analytics for incomplete PMU measurements-based power system stability assessment," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 1124–1126, Jan. 2018.

[25] C. Ren and Y. Xu, "A fully data-driven method based on generative adversarial networks for power system dynamic security assessment with

missing data," *IEEE Transactions on Power Systems*, vol. 34, no. 6, pp. 5044–5052, Nov. 2019.
[26] J. J. Gu, H. Y. Chen, G. L. Liu, G. Q. Liang, X. L. Wang, and J. H. Zhao, "Super-resolution perception for industrial sensor data," arXiv preprint arXiv: 1809.06687.
[27] K. M. He, X. Y. Zhang, S. Q. Ren, and J. Sun. "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 770–778.
[28] C. L. P. Chen and Z. L. Liu, "Broad learning system: An effective and efficient incremental learning system without the need for deep architecture," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 1, pp. 10–24, Jan. 2018.
[29] Siemens Power Technologies International, "*PSSE 33.0 Program Application Guide: Volume-II*," Mar. 2013.
[30] R. Zhang, Y. Xu, Z. Y. Dong, and K. P. Wong, "Measurement-based dynamic load modelling using time-domain simulation and parallel-evolutionary search," *IET Generation, Transmission & Distribution*, vol. 10, no. 15, pp. 3893–3900, Nov. 2016.
[31] E. G. Potamianakis and C. D. Vournas, "Short-term voltage instability: effects on synchronous and induction machines," *IEEE Transactions on Power Systems*, vol. 21, no. 2, pp. 791–798, May 2006.
[32] E. Keogh and C. A. Ratanamahatana, "Exact indexing of dynamic time warping," *Knowledge and Information Systems*, vol. 7, no. 3, pp. 358–386, Mar. 2005.
[33] J. S. Armstrong and F. Collopy, "Error measures for generalizing about forecasting methods: Empirical comparisons," *International Journal of Forecasting*, vol. 8, no. 1, pp. 69–80, Jun. 1992.
[34] *IEEE Standard for Synchrophasor Data Transfer for Power Systems*, IEEE Standard C37.118.2–2011, Dec. 2011, pp. 1–53.

Chao Ren received the B.E. degree from the School of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2017. He is currently working toward the Ph.D. degree in Cross-disciplinary of Computer Science and Electrical Engineering from the Interdisciplinary Graduate School, Nanyang Technological University, Singapore. He won several programming contest awards, including the champion of Chinese Software Cup, etc. His research interests include machine learning, data-analytics, optimization, security assessment and its applications to power engineering.

Yan Xu (S'10–M'13–SM'19) received the B.E. and M.E degrees in Electrical Engineering from South China University of Technology, Guangzhou, China in 2008 and 2011, respectively, and the Ph.D. degree in Electrical Engineering from The University of Newcastle, NSW, Australia, in 2013, all in electrical engineering. He is currently an Associate Professor at School of Electrical and Electronic Engineering, Nanyang Technological University (NTU), and a Cluster Director at Energy Research Institute @ NTU (ERI@N), Singapore. Previously, he held The University of Sydney Postdoctoral Fellowship in Australia. His research interests include power system stability and control, microgrid, and data-analytics for smart grid applications. Dr Xu is an Editor for IEEE TRANSACTIONS ON SMART GRID, IEEE TRANSACTIONS ON POWER SYSTEMS, CSEE Journal of Power and Energy Systems, and an Associate Editor for IET Generation, Transmission & Distribution.

Junhua Zhao (M'07–SM'17) received the Ph.D. degree from The University of Queensland, St Lucia, QLD, Australia. He is the External Experts of "Australian National Outlook", the Co-chair of the IEEE Special Interest Group (SiG) on Active Distribution Grids and Microgrids as well as the Secretary of the Asia Pacific Working Group of the IEEE PES SBLC (Smart Building, Load and Customer). He is a member of a Global Smart Grid Federation (GSGF) working group. Besides, he is a member of the "Interfaces of Grid Users/Focus on EV and Local Storage" working group, Smart Grid Australia (SGA) working group, "Cyber Physical Security of the Smart Grid" group and "Critical Infrastructure Program for Modelling and Analysis (CIPMA)" expert group. He is the Vice-chair of the Shenzhen AI Industry Society's Expert Committee. He is the editorial board members of IET Energy Conversion and Economics, Journal of Modern Power Systems and Clean Energy, Electric Power Components and Systems, and Power System Protection and Control. He is the expert reviewer of Australian Research Council (ARC), National Natural Science Foundation of China Reviewer, and Hong Kong Research Grants Committee (RGC). Also, he is the reviewer of IEEE Transactions on Power Systems, IEEE Transactions on Smart Grid, IEEE Transactions on Neural Networks and Learning Systems, Applied Energy, and IET Generation, Transmission & Distribution. His research interests include smart grid, electricity market, energy economic, data mining, and artificial intelligence.

Rui Zhang received the B.E. degree from the University of Queensland, and the Ph.D. degree from the University of Newcastle, Australia, in 2009 and 2014, respectively. She is now with University of New South Wales, Australia, and also a guest professor at Changsha University of Science and Technology, China. Her research interests include power system operation, control, and stability.

Tong Wan obtained BEco Honours from Queensland University of Technology, Australia and MPhil (Electrical Engineering) from the University of Sydney in 2007 and 2021 respectively. Her research interest is energy storage and renewable energy planning and risk management.