International Journal Of Industrial Electronics Control and Optimization
Print ISSN: 3517-2645
Online ISSN: 3568-2645
Homepage: https://ieco.usb.ac.ir/

Multi-class short-term voltage stability assessment considering the missing data

Amir Hossein Babaali[1] | Mohammad Taghi Ameli[2]

Faculty of Electrical Engineering, Shahid Beheshti University, Tehran, Iran[1,2]
Corresponding author's email: mtameli@yahoo.com

Article Info

Article type:
Research Article

Article history:
Received: *************
Received in revised form: *************
Accepted: *************
Published online: ************

Keywords: gated recurrent unit, communication delay, missing data, short-term voltage stability, multi-class classification.

ABSTRACT

Short-term voltage stability (STVS) varies with operating conditions of power networks, making its accurate assessment a critical challenge. This paper investigates a multi-class, data-driven approach to STVS assessment. A dynamic index is employed to categorize voltage magnitude variations into three classes: stable, alert, and unstable. A significant obstacle in data-driven methods is missing measurement data, typically caused by sensor failures or communication delays. To address this issue, we propose two complementary solutions. First, a Bidirectional Gated Recurrent Unit (Bi-GRU) network with an attention mechanism is designed to recover data loss due to sensor failures. This method leverages both temporal trends and historical system information to reconstruct missing values with high accuracy. Second, a variable-length sliding window (VLSW) algorithm combined with a Bi-GRU is introduced to mitigate data loss arising from communication delays. The VLSW algorithm enhances data diversity and enables fast recovery. Simulation results on IEEE 39-bus and IEEE 118-bus test systems demonstrate that the proposed framework effectively identifies multi-class STVSA under missing data conditions and remains robust against long-range data losses. Finally, validation on a real-world local network further confirms the practicality and robustness of the proposed approach.

I. Introduction

A. Motivation

Voltage stability has emerged as a critical concern in modern power systems characterized by high penetration of renewable energy sources (RESs) and the proliferation of microgrids. Unlike traditional synchronous generators, inverter-based RESs—such as solar photovoltaic (PV) and wind turbines—lack inherent inertia and reactive power support, compromising the system's ability to regulate voltage and maintain dynamic stability [1],[2]. Microgrids, capable of operating either connected to or islanded from the main grid, introduce additional complexity due to their limited voltage control mechanisms and heightened vulnerability to disturbances [3].

In this context, data-driven methods have gained considerable traction for real-time monitoring, stability assessment, and control, leveraging the widespread deployment of Phasor Measurement Units (PMUs) and advanced measurement infrastructure [4]. However, a persistent and significant challenge in these methods is missing data, which may arise from sensor malfunctions, communication failures, latency issues, or cyber-physical disruptions. Missing PMU measurements can severely impair the accuracy of voltage stability assessment models—particularly those based on dynamic, multi-class classification approaches [5].

The quality of PMU data severely restricts its various applications in the power system, in which a major factor affecting data quality is PMU data loss [6]. Due to communication congestion, hardware failures, data transmission/communication delays, cyber-attacks, and GPS-time synchronization issues, on-site PMUs typically experience different degrees of data quality issues[7]. Whenever data loss happens, the data-driven process of short-term voltage stability assessment (STVSA) cannot work correctly. Therefore, this paper focuses on the issue of STVSA considering missing data.

B. Literature Review

Many research efforts have been made to evaluate STVS online based on measurement data and machine learning (ML) algorithms Some promising ML-based data-driven solutions have been recently reported in the literature [8],[9]. Since STVS is generally driven by various fast-acting dynamic loads, such as induction motors and power electronically controlled devices[10], it often presents complicated stability patterns and characteristics in practical complex power systems.

Recently, deep learning-based neural network methods [11], [12],[13], [14], [15], [16] have been used with significant advantages in data-driven STVS assessment. The spatial-temporal features have been employed to learn a recurrent neural network (RNN) for STVSA[11], and the combination of graph convolutional network (GCN) and RNN is presented in [12]. Spatial-temporal GCN (STGCN) proposes to assess online STVS [13]. STGCN replaces the RNN module with the parallel-enabled 1-D Convolutional neural network (CNN). With the introduction of the attention mechanism, a gated recurrent graph attention network is utilized in [14] to perform spatial-temporal correlation learning, which results in adaptive STVSA against topological changes. A generative adversarial network (GAN) -based data augmentation in [15] and a data-driven approach to cyberattack conditions in [16] is presented to solve small data and bad data in STVS assessment.

Despite the efforts made, a serious gap is seen in the studies. Concretely [9],[12],[13] assumed that the measurement data are completely and continuously available. Some studies like [11] have investigated the impact of data loss but have not provided a solution.

To address this, data-driven approaches combining recovery and classification have been widely explored. For example, Long short-term memory (LSTM)-based imputation with Double Deep Q-Learning (DDQN) has been used to reconstruct missing data and classify voltage stability states under moderate data loss [17], while adaptive transient stability assessment frameworks explicitly considering PMU failures have demonstrated improved accuracy compared to mean imputation and tree-based classifiers [18]. A multi-task learning models [19] have been introduced to enhance robustness against incomplete or noisy measurements.

Ref.[20] presents an ensemble representation learning scheme (ERLS) to achieve data loss-tolerant online STVS assessment. It has focused on temporal missing data to recover lost data in different conditions. Attempts have been made to address the data loss issue in power system dynamic security assessment (DSA) [21]. These works focus mostly on network observability and cannot recover temporally lost data.

In practice, the power system operates at different operating points. Therefore, voltage stability should be monitored in several categories to create control measures appropriate to each operating situation. Some states may be alert but in binary classification, it recognizes unstable. So, the network can be brought to a stable state with limited instructions and low cost. Contrary to this insight, many studies including [13]-[20]. have assessed STVS in a binary class, which is a weakness. This paper attempts to fill this gap by examining multi-class STVS and modeling missing data in this framework.

PMU data recovery methods fall into two groups. The first uses correlations among PMUs, such as state estimation [22], tensor methods with historical data [23],[24], or correlation-based recovery[25]. While effective, these approaches degrade when multiple PMUs fail and tensor methods are too costly for real-time use; some variants also exploit past data [26].

The second group is prediction-based, relying mainly on historical PMU records [27], [28]. For example, [27] combines temporal patterns with spatial correlations from neighboring PMUs. These methods can handle wide data loss, but during sudden system events, short histories alone cannot capture future dynamics.

C. Innovations

A review of previous studies reveals that many studies ignored missing data. Moreover, most studies treated voltage stability as a binary problem, despite the fact that real measurement data are often affected by noise, communication interruptions, or PMU failures. Relying solely on a binary stable/unstable signal is insufficient for network monitoring and timely corrective actions. In contrast, this paper evaluates online voltage stability in a multi-class framework while explicitly considering missing data.

In this study, we address multi-class STVSA with missing measurements. A dynamic index is proposed to classify voltage stability levels, and a Bi-GRU network is developed to recover missing data and perform data-driven STVS assessment. Missing data caused by PMU failures and communication delays are studied in detail, and solutions are presented to reconstruct the data quickly and accurately.

Based on a thorough literature review and the limitations of existing approaches, the main contributions of this work are:

(1) Investigation of multi-class STVS under missing data conditions, accompanied by a dynamic index for accurate voltage stability classification.

(2) Integration of PMU data recovery with recent advances in artificial neural networks by developing a Bi-GRU network for both data reconstruction and voltage stability assessment.

(3) Detailed consideration of two types of missing data—PMU failure and communication delays—and presentation of solutions to ensure high-accuracy recovery.

(4) Analysis of the impact of recovered data on the accuracy of voltage stability predictions.
(5) Validation of the proposed model on a real power system, demonstrating its practical applicability.

The remainder of the paper is organized as follows: the proposed methodology is first introduced, followed by a description of the mathematical model, and finally, simulation results and discussions are presented.

II. Proposed model

This study aims to determine STVS based on measured data. The measurement data is considered in a time series format. Based on a dynamic index and behavior of the voltage magnitude, the database is classified into three classes: stable, alert, and unstable.

Measurement data may not be available due to PMU failures or communication delays during operation. The other part of the work deals with missing data. Measurement data is assumed to be unavailable for several seconds (PMU failure) or operating cycles (communication delays).

Time series data shows the dynamic behavior of the system well. One of the deep learning methods for data augmentation is a recurrent neural network (RNN), which is trained to process and convert a sequential data input into a specific sequential data output. GRU is an improved model of RNN, which can learn the dependency of past and future data simultaneously and thus predict future data with high accuracy. Among the recurrent neural networks, GRU has a simpler but more efficient model for the data of this paper, so a GRU model is developed for this study.

Fig. 1 shows the proposed framework for STVSA under missing data. On this basis, the lack of access to the data cannot be a problem in voltage stability assessment. Once access to the PMU data is interrupted, alternative data is predicted based on the behavior of previous data and the data sequence to determine the voltage stability.

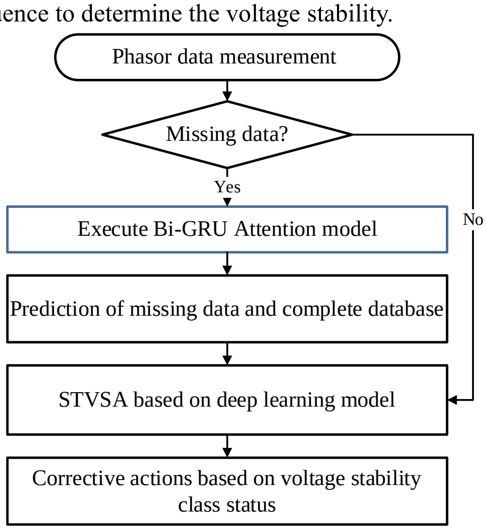

Fig. 1. STVSA with missing data in an overview

A. Failure of PMUs

One of the cases of data loss is the failure or out-of-service of PMUs. Fig. 2 shows that predicting missing data is based on a Bi-GRU and Attention layer learning process. In this structure, a masking layer is considered to remove null vectors in the input sequences. The Bi-GRU is compared in forward and backward forms.

A unidirectional GRU is included in the decoder stage. The dropout layer is used in both the encoder and decoder layers to avoid data overfitting. A dense layer is defined with a linear activation function to generate projections with continuous values. Missing data is retrieved using the previous measurement data through Bi-GRU in two stages: encoder and decoder. Past data samples are analyzed in the encoding phase.

As the data retrieval and training process progresses, decoding is performed, and new data is predicted. At this stage, the attention mechanism helps to focus on valuable data to increase the accuracy of data prediction.

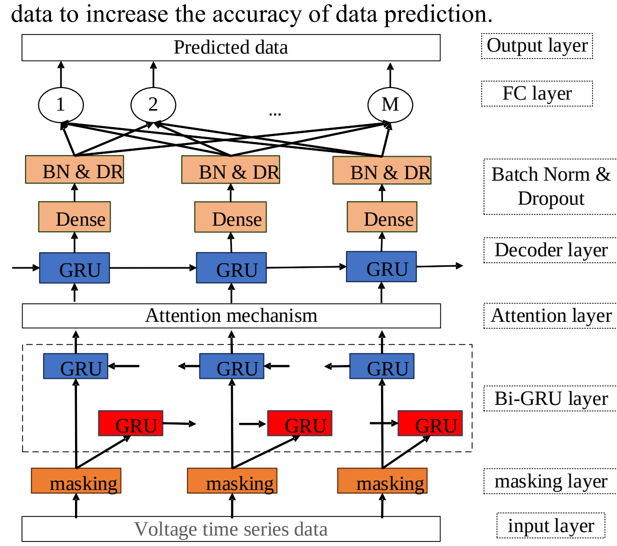

Fig. 2. The framework of the proposed method for recovering missing data in STVSA

B. Communication delay problem

We propose an attention-Bi-GRU approach to recover missing data caused by communication delays (see Fig. 2). In this section, we suggest the Variable-Length Sliding Window (VLSW) algorithm to recover limited data more accurately. The VLSW algorithm considers both past and future time indicators. Besides, it can generate a significantly greater number of samples by dynamically altering the input sequence length during the sampling process, which is crucial for implementing deep learning models.

The structure of Fig. 2 is based on the sequence-to-sequence method, where the encoder and decoder are two key functions. The encoder processes an input time series and maps it to a high-dimensional vector. The decoder takes the input from the vector and calculates the target data sequence. The attention mechanism also allows the decoder to learn how to focus on a particular region of the input sequence for different outputs.

III. Mathematical model

As **Error! Reference source not found.** shows that the recovery of the missing data is defined based on a Bi-GRU, whose model is presented in this section. The input for the problem is a sequence of variable length $x=\{x_1,x_2,...,x_n\}$, where *d* represents the features at a time index i. At each time index, the GRU maintains the hidden state h that leads to the hidden sequence $h=\{h_1,h_2,...,h_n\}$. The hidden vector at time index t is updated as follows.

$$i_t=\sigma(W_{xi}x_t+W_{hi}h_{t-1}+b_i) \quad (1)$$

$$f_t=\sigma(W_{xf}x_t+W_{hf}h_{t-1}+b_f) \quad (2)$$

$$c_t=f_t\otimes c_{t-1}+i_t\otimes\tanh(W_{xc}x_t+W_{hc}h_{t-1}+b_c) \quad (3)$$

$$O_t=\sigma(W_{xo}x_t+W_{ho}h_{t-1}+b_o) \quad (4)$$

$$h_t=o_t\otimes\tanh(c_t) \quad (5)$$

In the above relationships, i, *f* and *o* represent input, forgetting gates, and output. *c*, σ, and ⊗ are the cell vector, the sigmoid function, and the element-wise multiplication, respectively. In Bi-GRU, the input sequences are read not only from one sequence of h but from both the forward $\vec{h}=\{\overrightarrow{h_1},\overrightarrow{h_2},\dots,\overrightarrow{h_n}\}$ and backward $\overleftarrow{h}=\{\overleftarrow{h_1},\overleftarrow{h_2},\dots,\overleftarrow{h_n}\}$ directions to extract two sequences with hidden states, and then the hidden states are integrated into the time index i as follows:

$$h_i=[\overrightarrow{h_i},\overleftarrow{h_i}] \quad (6)$$

The output of the Bi-GRU encoder is a sequence of hidden states of length *n* with input sequence *x*.

A. GRU decoder with an attention mechanism

The output recursive sequence $\mathrm{y}=\{\mathrm{y}_1,\mathrm{y}_2,\mathrm{y}_3,\dots,\mathrm{y}_\mathrm{m}\}$ is generated by the decoder. In this study, the GRU decoder is proposed with an attention mechanism similar to the attention structure in [29]. The attention mechanism allows the model to pay attention to certain parts of the data and give more weight to these parts when making predictions.

One issue with this model is the assignment of continuous values as output. Instead of calculating the probability distribution for the following possible outputs, a fully connected layer with a linear activation function is added to the GRU (Gated Recurrent Unit) layer. This modification allows the model to produce predictions in continuous values, which are used to estimate missing data. For the decoder, the estimate at time t is calculated as follows:

$$y_t=Linear(W\,[s_t;c_t]+b) \quad (7)$$

$$s_t=GRU\,(y_{t-1},s_{t-1},c_t) \quad (8)$$

where s_t is the hidden state of the decoder at time index t, c_t is the attention context vector (for each output y_t) and $[s_t,c_t]$ is a combination of the decoder's hidden state and the context vector. The parameters W and b map this combination to the size of the hidden states of the decoder and generate the final linear prediction layers y_t.

At each time sequence t, the attention context vector can be described as the weighted sum of the hidden states transmitted by the Bi-GRU encoder:

$$c_t=\sum_{i=1}^{n}\alpha_{ti}h_i \quad (1)$$

The weight of each hidden state α_{ti} is calculated as follows:

$$e_{ti}=a(s_{t-1},h_i), \quad (2)$$

$$\alpha_{ti}=\mathrm{softmax}() \quad (3)$$

where e_{ti} expresses the dependency between the hidden states surrounding the h_i and the output in the sequence *t*. *a* is a feedforward network that can be jointly trained with other components of the GRU decoder. To ensure that the sum of all attention weights is normalized to 1, the softmax activation function is applied to e_{ti}.

As seen in Fig. 2, the masking layer is embedded after the input layer. The input layer includes a sequence of time series. The masking layer filters the vectors with a null layer in the input samples according to the input data and sends them to the Bi-GRU encoder. The encoder processes the data in future and past time sequences and generates a hidden output vector. Then the GRU decoder processes the hidden vector and the dense layer recursively generates the final target sequence. During this process, the decoder calculates the attention weights in each time sequence to focus on specific input parts and extract relevant information.

B. Variable length sliding window (VLSW) algorithm

During the data-driven STVSA process, the measurement data may arrive at the operation center with a delay or in a discrete form. In this section, assuming a communication delay in the measurement data, the variable length sliding window (VLSW) algorithm is used along with the proposed method (see Fig. 2). This way, alternative samples are produced quickly with only a few available samples.

One advantage of this method is the simultaneous use of past and future time sequences. In addition, dynamically changing the input sequence length in the sampling process generates more samples, which improves machine learning performance.

Fig. 3 shows how the VLSW algorithm works, using a dynamically changing range of input data to generate prediction patterns. In this algorithm, both the length of the left time series ($|IL|$) and the length of the right time series ($|IR|$) change during sampling. Assuming that the $|IL|$ varies dynamically between *m* and *n* and $|IR|$ varies dynamically between *o* and *p*, the sample number generated is Eq.(4).

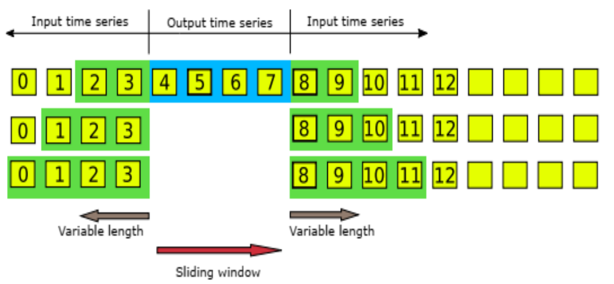

Fig. 3. Illustration of the VLSW algorithm

VLSW can generate more samples than the sliding window algorithm with future information [30], which can lead to higher accuracy in predicting missing data.

$$N=\sum_{i=m}^{n}\sum_{j=o}^{p}|T|-i-|O|-j+1$$
$$=\frac{(p-o+1)(n-m+1)\left[2(|T|-|O|+1)-(m+n+o+p)\right]}{2} \quad (4)$$
$$, n+p+|O|\le|T|$$

C. Model evaluation

Root Mean Square Error (RMSE), and Mean Absolute Error (MAE) are methods to evaluate the performance of missing data augmentation. The mathematical relationship between these methods is given below [7].

$$RMSE=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\left|f_i-\hat{f}_i\right|\right)^2} \quad (5)$$

$$MAE=\frac{1}{n}\sum_{i=1}^{n}\left|f_i-\hat{f}_i\right| \quad (6)$$

Where f_i is the real value and $\hat{f}_i$ is the estimated value of the monitored parameters. We use RMSE and MAE criteria because these can show the sensitivity of the proposed model to outlier data.

IV. Simulations and results

The investigated networks are IEEE 39-bus and 118-bus networks. The 39-bus network is used to recover the missing data caused by the PMU failure, and the 118-bus network is used to recover the missing data caused by the communication delay. According to [31], IEEE 39-bus and 118-bus have 8 and 21 buses connected to the PMU, respectively (TABLE I) to maintain the observability of the network.

According to the frequency of the 39-bus network (60 Hz) and the 118-bus network (50 Hz), the data sampling rates are assumed to be 30 ms and 20 ms, respectively. A time domain simulation is performed to generate operation scenarios and create a database in the DIgSILENT Power Factory 15.1.7-application environment and RMS mode [32]. The database is classified according to dynamic indices [32] in the MATLAB 2019b environment, and the learning and evaluation process is implemented in Python.

TABLE I. Selected buses connected to PMU in case studies

Case study	Selected buses (connected to PMU)
IEEE 39bus	3,8,10,16,20,23,25,29
IEEE 118bus	3,13,29,35,43,47,55,58,62,63,69,75,81,82,9 1,93,104,106,107,113,115

A. Classification of voltage time series data

In this study, the dynamic behavior of voltage is considered for the STVS assessment, so a database has been created based on time series. Static indicators to determine voltage stability only work from a certain operating point, so they are unsuitable for classifying voltage time series. To classify time series data, a dynamic index is needed to include all voltage change characteristics. Here, a dynamic voltage index is defined by considering the NERC operating voltage guidelines and the Lyapunov exponent (LE) [32].

a. Lyapunov exponent

LE is a fundamental concept of nonlinear dynamics that quantifies the local stability features of attractors and other invariant sets in state space. Positive LE indicates the exponential divergence of neighboring trajectories and is critical of chaotic attractors [32]. Assume a continuous dynamic system: $\dot{x}=f(x), x\in X\subset R^N$. If $\phi(t,x)$ is the answer to differential equations, the following restricted matrix is defined:

$$\Lambda(x)=\lim_{t\to\infty}\left[\frac{\partial\Phi(t,x)}{\partial x}^T\frac{\partial\Phi(t,x)}{\partial x}\right]^{1/2t} \quad (7)$$

Eq.(8) shows the LE per bus i for changes in the voltage magnitude. So that $\Lambda_i(x)$ is the eigenvalues of the bounded matrix $\Lambda(x)$. The LE ($\lambda_i(x)$) is defined as $\lambda_i(x)=log\Lambda_i(x)$.

In Eq.(8), $k=1,2,\ldots,M$. Δt, M and N_b are the sampling rate, time window, and number of busses, respectively. For fixed and small values $0<\varepsilon1<\varepsilon2$, N should be selected according to Eq.(8).

$$\lambda_i(k\Delta t):=\frac{1}{Nk\Delta t}\times\sum_{m=1}^{N}\log\frac{\left\|V_{(k+m)\Delta t}-V_{(k+m-1)\Delta t}\right\|}{\left\|V_{m\Delta t}-V_{(m-1)\Delta t}\right\|}, k>N \quad (8)$$
$$\forall i=1,2,\ldots,N_b$$

$$\varepsilon_1<\left\|V_{m\Delta t}-V_{(m-1)\Delta t}\right\|<\varepsilon_2, m=1,2,\ldots,N \quad (9)$$

For each bus, λ_i specifies the trend of the voltage curves. λ_i will be positive for divergent oscillatory trends; otherwise, it will be negative.

b. Voltage Deviation (VD)

According to NERC, a voltage recovery problem occurs when the voltage deviation remains above 0.2 for 20 cycles [32]. VD is calculated using the following equations.

$$\Delta V_i^t = \frac{V_i^0 - V_i^t}{V_i^0}, i \in [1, N_b], t \in [t_{cl}, t_a] \quad (10)$$

$$VD_i = \min\left\{\Delta V_i^t\right\} \quad (11)$$

In Eq.(10), the voltage deviation per bus is calculated from the pre-fault value or steady-state, where V_i^0 is the pre-fault voltage in the i-th bus and t_a is equal to 20 cycles of the nominal frequency after clearing time (t_{cl}).

The voltage stability status is classified by combining VD and LE (or λ) into stable, alert, and unstable classes, according to TABLE II. If the oscillating behavior of the voltage is divergent (it means λ>0) and the voltage does not return within the defined time (VD>0.2), voltage instability has occurred (The red image in Fig. 4). In the alert mode, the voltage returned to an acceptable range after the fault clearing (VD<0.2), but voltage oscillating behavior resulted in a positive LE (The yellow image in Fig. 4). Moreover, a stable mode is shown by the green image in Fig. 4.

TABLE II. Database classification summary based on dynamic index and Lyapunov exponent

Lyapunov exponent (λ)	Voltage deviation (VD)	Stability class
0<	0.2<	Unstable
0>	0.2<	Alert
0<	0.2>	Alert
0>	0.2>	Stable

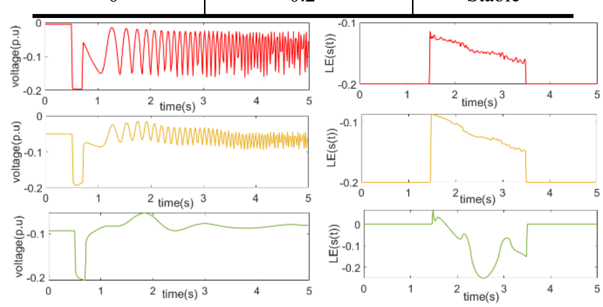

Fig. 4. Voltage magnitude curve and Lyapunov exponent in three-class STVS

B. Missing data due to PMU failure

To evaluate the performance of the proposed model in the 39-bus network (see Fig. 5).

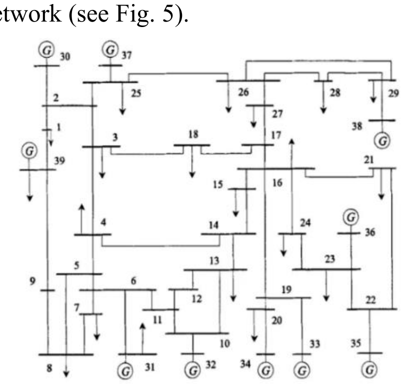

Fig. 5. One-line diagram of IEEE 39-bus network

We developed 1,000 operating scenarios that simulated three-phase faults occurring on the lines of the 39-bus system and generating unit failures. A database is created through time-domain simulations and classified into three classes: stable, alert, and unstable. This database is divided into training data (70%) and test data (30%). The parameters for the attention-Bi-GRU model are adjusted as detailed in TABLE III. Attention size represents the linear size of the attention weights.

TABLE III. Setting the main parameters of the attention-Bi-GRU model

Parameters	quantity/ type
Activator function	Tanh and ReLU
Learning rate	10e-5
Epochs	200
Batch size	128
Dropout	0.3
Attention size	16
hidden layer unit	128

Assuming we cannot access any of the PMUs, its data is predicted using the proposed method (Attention-Bi-GRU). The predicted data is compared with the actual data to evaluate the model, and MAE and RMSE criteria are formed. The results for a set of PMUs are shown in TABLE IV. It includes simulation results of the proposed method (Attention-Bi-GRU) and a generative adversarial network (GAN) with an attention mechanism.

In TABLE IV, voltage time-series is retrieved over 5 seconds. According to the results obtained, the proposed method predicted the missing data with higher accuracy than the GAN-attention method, so that for all PMUs, RMSE and MAE are less than 0.18 % and 0.16 %, respectively.

TABLE IV. The function of predicting missing data in each of the PMUs through the proposed model and GAN-attention model

PMU No.	**RMSE**		**MAE**	
	Att-Bi-GRU	GAN-Att	Att-Bi-GRU	GAN-Att
PMU#3	1.78E-03	1.86E-03	1.61E-03	1.74E-03
PMU#10	1.48E-03	1.59E-03	1.40E-03	1.44E-03
PMU#20	1.67E-03	1.78E-03	1.59E-03	1.69E-03
PMU#25	1.55E-03	1.68E-03	1.51E-03	1.60E-03
PMU#29	1.64E-03	1.76E-03	1.56E-03	1.66E-03

This paper proposes an approach that simultaneously recovers lost data and evaluates STVS. We compare our method with several state-of-the-art methods from various sources. Fig. 6 illustrates that the alternative data through the attention Bi-GRU resulted in a high accuracy of about 99% in the STVS assessment.

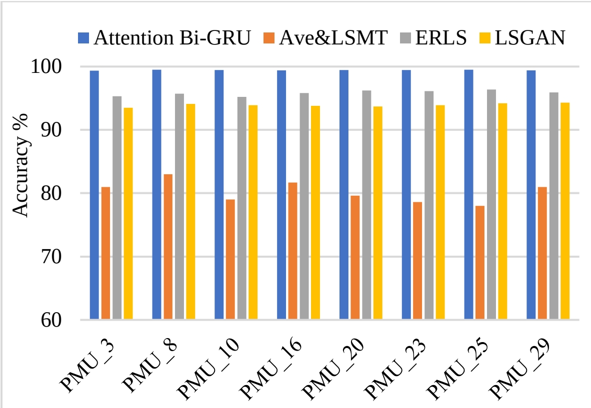

Fig. 6. Comparison of STVSA with recovering missing data due to PMU failure via various methods

As illustrated in Fig. 6, the use of LSTM with historical averaging achieves only about 80% accuracy in STVSA. In comparison, the least squares GAN (LSGAN)-based augmentation [15] and ERLS [20] improve performance to 94% and 95.8%, respectively. The proposed Attention Bi-GRU further advances this result, reaching 99.4% accuracy, which highlights its stronger capability in handling missing data and ensuring reliable stability assessment.

a. Hyperparameter tuning

The sensitivity analysis in Fig. 7 indicates that increasing the attention size and hidden units enhances performance until a saturation point, beyond which improvements are negligible. For dropout, no regularization (0.0) leads to overfitting with unstable recovery errors, while a high rate (0.6) causes underfitting and accuracy degradation. A moderate dropout value (0.3), together with attention size 16 and 128 hidden units, achieves the most stable balance, yielding low RMSE/MAE and consistently high classification accuracy.

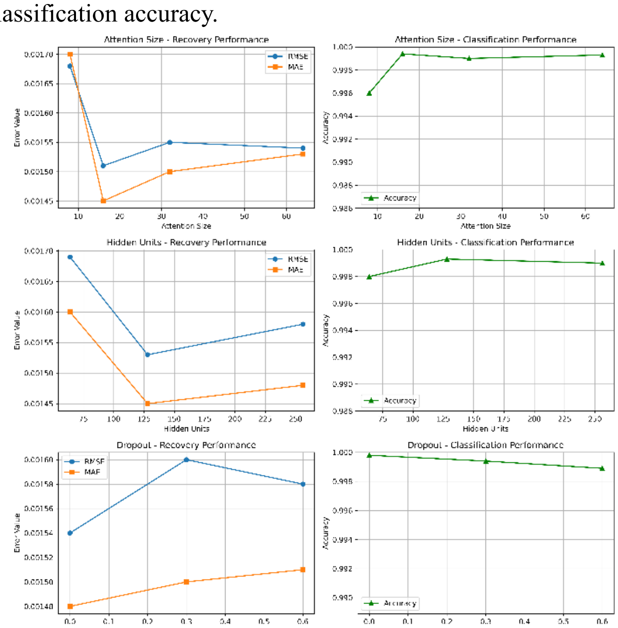

Fig. 7. Sensitivity analysis of key hyperparameters

C. Comparative Analysis with Existing Methods

To highlight the advantages of the proposed method, Fig. 8 provides a radar diagram comparing it with other ANN-based models. The proposed model demonstrates strong capability in recovering missing data and assessing multi-class STVS, whereas approaches such as LSGAN and ERLS fail to correctly identify multiple stability boundaries.

For further clarification, we assume the loss of PMU_3 data due to failure, and evaluate multi-class STVS using different methods across three categories: stable, alert, and unstable. Most alternative methods fail to detect the alert state effectively, mainly because their training datasets are limited to binary classification. Moreover, their assessment of stable and unstable conditions is inadequate, likely due to insufficient handling of missing data.

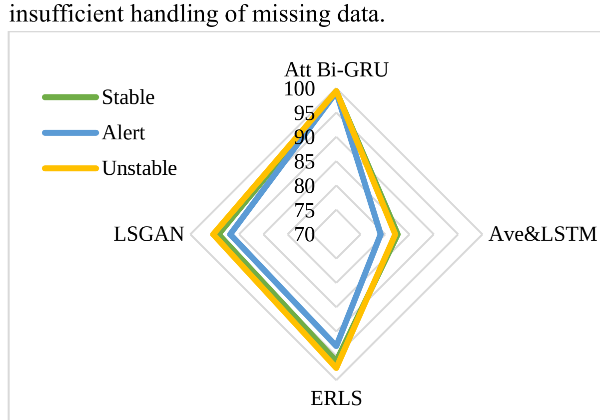

Fig. 8. Multi-class STVSA results through different methods with missing data in PMU_3

According to [33], this section considers 1000 operation scenarios in the IEEE 118-bus, including the failure of dense lines and the out-of-service of some power units, to generate the database. The database is divided into training and testing sets, in which training is based on the prediction model of Fig. 2.

D. Missing data due to communication delay

According to the IEEE C37.118.2-2011 standard, the communication delay between the PMU and PDC is typically between 0.02s and 0.05s [34]. The sampling rate consumed 20 ms in the IEEE 118-bus and the maximum permissible communication delay consumed 50 ms; three measurement data are missing for each delay. So, this study assumes three cases of communication delays with values of 50ms, 100ms, and 150ms, i.e., 3, 5, and 8 missing data samples. For each of these, the parameters setting of VLSW values are given in TABLE V.

In TABLE V, m, n, o, p, and q are the minimum and maximum length of the data before the missing data, the minimum and maximum length of the window after the missing data, and the output vector length, respectively, as shown in Fig. 3.

TABLE V. Setting the initial values of the parameters of the VLSW

parameters	50 ms		100 ms		150 ms		200 ms	
	w_1	w_2	w_1	w_2	w_1	w_2	w_1	w_2
M	3	3	3	3	3	3	4	4
N	5	5	6	6	8	8	9	9
O	3	3	3	3	3	3	4	4
P	5	5	6	6	8	8	9	9
q	3	5	5	8	8	10	10	11

In TABLE V, window 1 (w_1) shows that the length of the output vector at each delay is equal to the length of the missing data sequence. In w_2, it is assumed that the length of the output vector is longer, which means more data is lost.

Fig. 9-Fig. 11 present the performance of the proposed attention-based Bi-GRU model with VLSW across both training and test datasets. As shown in Fig. 9, the model successfully reconstructs missing data under a 50 ms communication delay (equivalent to three voltage magnitude samples at a 20 ms sampling rate). Despite the challenges introduced by this delay, the proposed framework achieves reliable convergence for both training and testing phases, reaching an accuracy of 98.5%.

The impact of increased communication delays is further demonstrated in Fig. 10, where the recovery accuracy decreases compared to Fig. 9. Specifically, with a longer delay leading to five consecutive missing data sequences, the reconstruction process becomes more complex. This results in reduced prediction speed and accuracy, highlighting that higher delays increase both the computational burden and the difficulty of maintaining high-fidelity recovery.

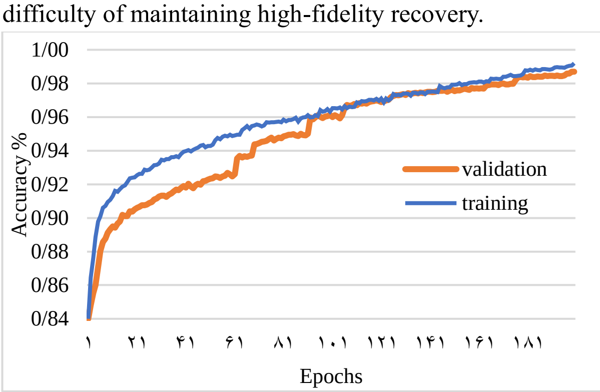

Fig. 9. The accuracy of training and testing the predicted data for a delay of 50 ms and w_1

Fig. 11 shows that the proposed model accurately predicted the alternative data, although the range of the missing data sequence increased. Thus, with the increased data required for prediction, the model's accuracy did not change significantly compared to a delay of 100 ms. The results indicate that attention Bi-GRU with the VLSW algorithm can accurately retrieve missing data due to communication delay. In this way, STVS can be assessed with complete information.

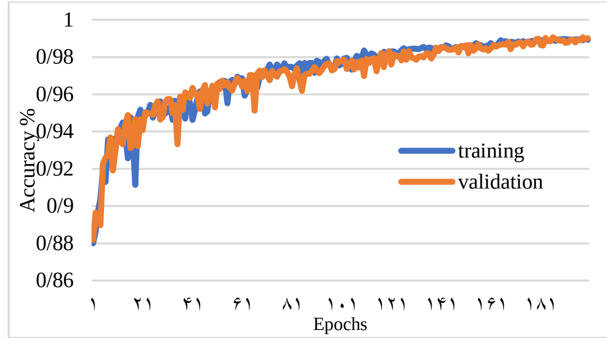

Fig. 10. The accuracy of training and testing the predicted data for a delay of 100 ms and w_1

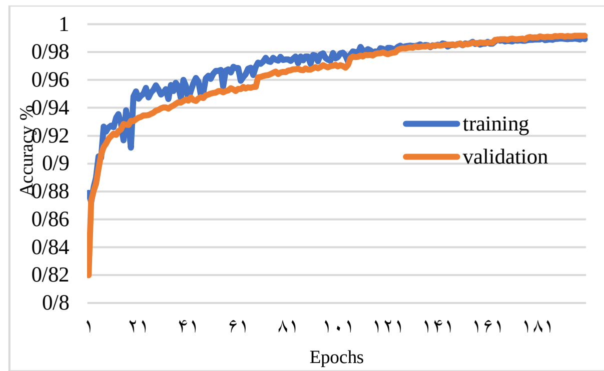

Fig. 11. The accuracy of training and testing the predicted data for a delay of 150 ms and w_1

Cases 1 to 3 represent communication delays of 50, 100, and 150 ms, respectively, while w1 and w2 correspond to different VLSW settings.

According to Fig. 12, the samples are recovered with an error of less than 0.2% across these scenarios, indicating that the estimated values remain very close to the actual missing data. Although the prediction error shows a slight upward trend as the delay increases, it consistently remains within an acceptable margin.

As illustrated in Fig. 12, when the delay is extended to Case 4 (200 ms), the error further increases, with RMSE reaching around 0.19 and MAE up to 0.184. This demonstrates that the model performance begins to degrade more noticeably at longer delays. Nevertheless, the overall error level remains relatively low, suggesting that the proposed approach is still reliable for delays up to 200 ms, beyond which the operational boundaries of the model may need to be carefully considered.

E. Robustness verification of method in noisy scenarios

To assess the robustness of the proposed framework, noisy PMU measurements with different signal-to-noise ratios (SNR = 20, 30, 40, and 50 dB) were evaluated, and the results are shown in Fig. 13. Even at severe noise levels

(SNR = 20 dB), the model achieved 99.05% accuracy, with RMSE and MAE remaining on the order of 10^{-3}.

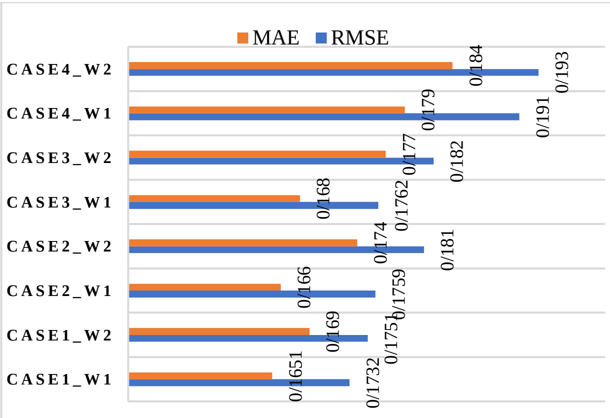

Fig. 12. Evaluation results of the proposed method for different modes of communication delay and sliding window

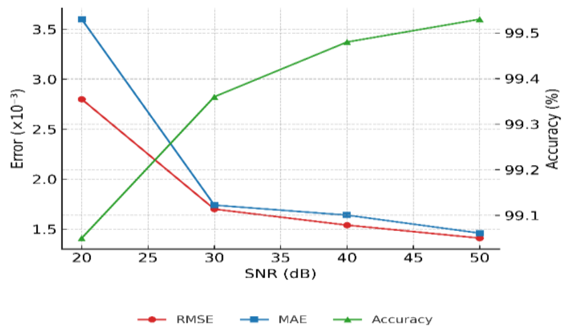

Fig. 13. Robustness of the proposed Attention Bi-GRU under different noise levels

As SNR increased, accuracy consistently exceeded 99%, while errors further decreased. These results demonstrate that the Attention Bi-GRU is not only effective for missing data recovery but also highly robust to measurement noise, confirming its practical applicability for real-world power grid monitoring.

F. Analysis of Computational Cost and Latency

In addition to accuracy and robustness, the computational efficiency of the proposed Attention Bi-GRU is evaluated to ensure suitability for real-time PMU applications. On a standard workstation (Intel i7 CPU, 16 GB RAM), the average inference time per PMU sequence is approximately 4 ms, which is significantly below the 20 ms reporting rate of conventional PMUs. The memory footprint remained modest (<250 MB during inference), and latency is negligible, confirming that the framework can operate under real-time constraints without imposing notable computational overhead. While training requires higher resources, this is performed offline and does not affect online deployment. These results indicate that the proposed model achieves a favorable trade-off between predictive accuracy and computational cost, ensuring practical scalability to large-scale power system monitoring.

G. Validation on Local network

A part of Iran Grid includes 72 buses, 47 transmission lines and 23 power units. We suppose 10 PMUs connected to 400 and 230 kV busses. Time domain simulations include 800 scenarios which consider N-1contingencies such as three-fault on lines and out-of-service of synchronous machines. TABLE VI shows average of RMSE and MAE as percentage in real power system. As shown, prediction of missing data in all of busses is done with high accuracy, so that, average of RMSE and MAE through proposed method (Att-Bi-GRU) are less than 0.14%.

TABLE VI. Results of missing data prediction in a Real

Method	RMSE	MAE
GAN-Att	0.174	0.171
Att-Bi-GRU	0.139	0.136

V. Conclusions

This paper introduces a data-driven framework for multi-class STVSA under conditions of missing data. By employing a dynamic index, STVS is categorized into three classes—stable, alert, and unstable—offering a more granular perspective compared to traditional binary approaches. To address the critical issue of incomplete measurements, two complementary strategies are proposed: (i) an attention-based Bi-GRU network for recovering data lost due to PMU failures, and (ii) a variable-length sliding window (VLSW) integrated with Bi-GRU to handle data loss caused by communication delays.

The proposed Attention Bi-GRU framework has demonstrated superior accuracy (over 99%) in short-term voltage stability assessment, even under challenging conditions of missing or noisy PMU data. Moreover, the method remains robust against severe communication delays (up to 200 ms) and noisy measurements (SNR as low as 20 dB), while maintaining RMSE and MAE in the order of 10^{-3}. These findings confirm the model's strong potential for real-time deployment in power system monitoring and control.

Validation on a real power system further confirmed its practicality, with both RMSE and MAE remaining under 0.14%. In addition to superior accuracy, the proposed framework benefits from reduced model complexity and faster operator response, making it suitable for real-time power system operations.

Recent studies have introduced multi-class stability assessment methods based on CNN-LSTM hybrids and transformer architectures. Although such approaches are not included in our benchmark analysis due to scope limitations, the superior accuracy of the proposed Attention Bi-GRU highlights its strong competitiveness with these state-of-the-art models.

Nevertheless, several limitations should be acknowledged. First, the framework relies on PMU placement quality, and suboptimal deployment may affect recovery performance. Second, although the model shows robustness to random noise and missing data, its resilience to adversarial or cyber-induced data losses was not explicitly investigated. Third, the threshold values for Lyapunov exponent and voltage deviation were based on NERC guidelines, without dataset-specific optimization.

Future research may therefore explore:

(1) Integration of graph-based or physics-informed priors to guide data recovery.
(2) Evaluation under adversarial or malicious missing data patterns for cybersecurity assurance.
(3) Optimization of thresholds to maximize classification accuracy across diverse systems.
(4) Exploring probabilistic deep learning techniques (e.g., Bayesian RNNs, MC-Dropout) to provide uncertainty bounds and confidence intervals for both data recovery and stability assessment.

References

[1] X. Liang, H. Chai, and J. Ravishankar, “Analytical Methods of Voltage Stability in Renewable Dominated Power Systems: A Review,” Electricity, vol. 3, no. 1, pp. 75–107, Feb. 2022, doi: 10.3390/electricity3010006.

[2] P. Fakhrooeian, mehrdad Abedi, and peyman karimyan, “Optimal allocation and sizing of dynamic VAR support to improve short-term voltage stability considering wind farm and dynamic load model,” IECO, vol. 1, no. 1, June 2018, doi: 10.22111/ieco.2018.24388.1024.

[3] N. Khosravi, D. Çelik, H. Bevrani, and S. Echalih, “Microgrid stability: A comprehensive review of challenges, trends, and emerging solutions,” International Journal of Electrical Power & Energy Systems, vol. 170, p. 110829, Sept. 2025, doi: 10.1016/j.ijepes.2025.110829.

[4] J. Cao, M. Zhang, and Y. Li, “A Review of Data-Driven Short-Term Voltage Stability Assessment of Power Systems: Concept, Principle, and Challenges,” Mathematical Problems in Engineering, vol. 2021, pp. 1–12, Dec. 2021, doi: 10.1155/2021/5920244.

[5] F. Boronuosi, S. Azad, M.-T. Ameli, and M. R. Shadi, “Dynamic Security Assessment of Power Systems Using a Deep Learning and GAIN-Based Approach for Addressing Missing Data,” Results in Engineering, p. 106585, Aug. 2025, doi: 10.1016/j.rineng.2025.106585.

[6] A. Xue et al., “Method of amplitude data recovery in PMU measurements that considers synchronisation errors,” IET Generation Trans & Dist, vol. 14, no. 24, pp. 5746–5755, Dec. 2020, doi: 10.1049/iet-gtd.2020.0785.

[7] Y. Cheng, B. Foggo, K. Yamashita, and N. Yu, “Missing Value Replacement for PMU Data via Deep Learning Model With Magnitude Trend Decoupling,” IEEE Access, vol. 11, pp. 27450–27461, 2023, doi: 10.1109/ACCESS.2023.3254448.

[8] L. Zhu, C. Lu, I. Kamwa, and H. Zeng, “Spatial–Temporal Feature Learning in Smart Grids: A Case Study on Short-Term Voltage Stability Assessment,” IEEE Transactions on Industrial Informatics, vol. 16, no. 3, pp. 1470–1482, 2020, doi: 10.1109/TII.2018.2873605.

[9] C. Ren, Y. Xu, Y. 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, June 2020, doi: 10.1109/TII.2019.2940098.

[10] J. Zhang, Y. Luo, B. Wang, C. Lu, J. Si, and J. Song, “Deep Reinforcement Learning for Load Shedding Against Short-Term Voltage Instability in Large Power Systems,” IEEE Transactions on Neural Networks and Learning Systems, vol. 34, no. 8, pp. 4249–4260, Aug. 2023, doi: 10.1109/TNNLS.2021.3121757.

[11] L. Zhu, D. J. Hill, and C. Lu, “Intelligent Short-Term Voltage Stability Assessment via Spatial Attention Rectified RNN Learning,” IEEE Transactions on Industrial Informatics, vol. 17, no. 10, pp. 7005–7016, Oct. 2021, doi: 10.1109/TII.2020.3041300.

[12] G. Wang, Z. Zhang, Z. Bian, and Z. Xu, “A short-term voltage stability online prediction method based on graph convolutional networks and long short-term memory networks,” International Journal of Electrical Power & Energy Systems, vol. 127, pp. 106647–106647, 2021, doi: 10.1016/j.ijepes.2020.106647.

[13] Y. Luo, C. Lu, L. Zhu, and J. Song, “Data-driven short-term voltage stability assessment based on spatial-temporal graph convolutional network,” International Journal of Electrical Power & Energy Systems, vol. 130, pp. 106753–106753, Sept. 2021, doi: 10.1016/j.ijepes.2020.106753.

[14] H. Cai and D. J. Hill, “A data-driven distributed and easy-to-transfer method for short-term voltage stability assessment,” International Journal of Electrical Power & Energy Systems, vol. 139, pp. 107960–107960, 2022, doi: 10.1016/j.ijepes.2022.107960.

[15] Y. Li, M. Zhang, and C. Chen, “A Deep-Learning intelligent system incorporating data augmentation for Short-Term voltage stability assessment of power systems,” Applied Energy, vol. 308, pp. 118347–118347, 2022, doi: 10.1016/j.apenergy.2021.118347.

[16] C. Ren, X. Du, Y. Xu, Q. Song, Y. Liu, and R. Tan, “Vulnerability Analysis, Robustness Verification, and Mitigation Strategy for Machine Learning-Based Power System Stability Assessment Model Under Adversarial Examples,” IEEE Transactions on Smart Grid, vol. 13, no. 2, pp. 1622–1632, 2022, doi: 10.1109/TSG.2021.3133604.

[17] A. Liu, X. Guan, D. Sun, H. Jiang, C. Cui, and D. Fang, “Short Term Voltage Stability Assessment with Incomplete Data Based on Deep Reinforcement Learning in the Internet of Energy,” in Wireless Algorithms, Systems, and Applications, vol. 12938, Z. Liu, F. Wu, and S. K. Das, Eds., in Lecture Notes in Computer Science, vol. 12938. , Cham: Springer International Publishing, 2021, pp. 225–236. doi: 10.1007/978-3-030-86130-8_18.

[18] B. Ma, J. Yang, X. Peng, K. Jiang, D. Liu, and K. Cao, “An adaptive assessment method of power system transient stability considering PMU data loss,” IET Generation Trans & Dist, vol. 18, no. 24, pp. 4116–4133, Dec. 2024, doi: 10.1049/gtd2.13340.

[19] T. Luo and X. Jiang, “A novel multi-task learning method for evaluating short-term voltage stability with incomplete PMU measurements,” Complex Intell. Syst., vol. 10, no. 2, pp. 1971–1983, Apr. 2024, doi: 10.1007/s40747-023-01252-8.

[20] L. Zhu et al., “Robust Representation Learning for Power System Short-Term Voltage Stability Assessment Under Diverse Data Loss Conditions,” IEEE Transactions on Neural Networks and Learning Systems, vol. 35, no. 5, pp. 6035–6047, May 2024, doi: 10.1109/TNNLS.2023.3325542.

[21] B. Tan et al., “Spatial-temporal adaptive transient stability assessment for power system under missing data,” International Journal of Electrical Power & Energy Systems, vol. 123, pp. 106237–106237, 2020, doi: 10.1016/j.ijepes.2020.106237.

[22] K. Sun, M. Huang, Z. Wei, and G. Sun, “High-Refresh-Rate Robust State Estimation Based on Recursive Correction for Large-Scale Power Systems,” IEEE Trans. Instrum. Meas., vol. 72, pp. 1–13, 2023, doi: 10.1109/TIM.2023.3277070.
[23] D. Osipov and J. H. Chow, “PMU Missing Data Recovery Using Tensor Decomposition,” IEEE Trans. Power Syst., vol. 35, no. 6, pp. 4554–4563, Nov. 2020, doi: 10.1109/TPWRS.2020.2991886.
[24] B. Foggo and N. Yu, “Online PMU Missing Value Replacement Via Event-Participation Decomposition,” IEEE Transactions on Power Systems, vol. 37, no. 1, pp. 488–496, Jan. 2022, doi: 10.1109/TPWRS.2021.3093521.
[25] Y. Cheng, B. Foggo, K. Yamashita, and N. Yu, “Missing Value Replacement for PMU Data via Deep Learning Model With Magnitude Trend Decoupling,” IEEE Access, vol. 11, pp. 27450–27461, 2023, doi: 10.1109/ACCESS.2023.3254448.
[26] Z. Yang, H. Liu, T. Bi, Z. Li, and Q. Yang, “An adaptive PMU missing data recovery method,” International Journal of Electrical Power & Energy Systems, vol. 116, pp. 105577–105577, 2020, doi: 10.1016/j.ijepes.2019.105577.
[27] J. Pei, Z. Wang, J. Wang, and D. Shi, “Robust fast PMU measurement recovery enhanced by randomized singular value and sequential Tucker decomposition,” IET Generation Trans & Dist, vol. 16, no. 16, pp. 3267–3281, Aug. 2022, doi: 10.1049/gtd2.12520.
[28] X. Deng et al., “Deep learning model to detect various synchrophasor data anomalies,” IET Generation Trans & Dist, vol. 14, no. 24, pp. 5739–5745, Dec. 2020, doi: 10.1049/iet-gtd.2020.0526.
[29] Z. Niu, G. Zhong, and H. Yu, “A review on the attention mechanism of deep learning,” Neurocomputing, vol. 452, pp. 48–62, Sept. 2021, doi: 10.1016/j.neucom.2021.03.091.
[30] Y.-F. Zhang, P. J. Thorburn, W. Xiang, and P. Fitch, “SSIM—A Deep Learning Approach for Recovering Missing Time Series Sensor Data,” IEEE Internet of Things Journal, vol. 6, no. 4, pp. 6618–6628, Aug. 2019, doi: 10.1109/JIOT.2019.2909038.
[31] M. M. Ahmed, M. Amjad, M. A. Qureshi, K. Imran, Z. M. Haider, and M. O. Khan, “A Critical Review of State-of-the-Art Optimal PMU Placement Techniques,” Energies, vol. 15, no. 6, p. 2125, Mar. 2022, doi: 10.3390/en15062125.
[32] A. H. Babaali and M. T. Ameli, “Weighted ensemble learning for real-time short-term voltage stability assessment with phasor measurements data,” IET Generation, Transmission & Distribution, Mar. 2023, doi: 10.1049/gtd2.12809.
[33] A. Babaali and M. T. Ameli, “Short-term voltage stability prediction based on a Bidirectional Gated Recurrent Unit neural network using phasor measurement data in power systems,” Computational Intelligence in Electrical Engineering, vol. 15, no. 1, pp. 1–16, 2024, doi: 10.22108/isee.2023.135273.1585.
[34] S. Shah, S. Koley, and F. Malandra, “Experimental End-To-End Delay Analysis of LTE cat-M With High-Rate Synchrophasor Communications,” July 11, 2022, arXiv: arXiv:2207.04847. doi: 10.48550/arXiv.2207.04847.

Amir Hossein Babaali received his B.Sc. degree from the University of Kashan, Isfahan, Iran, in 2012 and his M.Sc. degree from the Department of Electrical Engineering, Shahid Beheshti University, Tehran, Iran, in 2014. He is a Ph.D. student in the Department of Electrical Engineering at Shahid Beheshti University. His research interests include power system simulation, operation, planning, and machine learning applications in power systems.

Mohammad Taghi Ameli received his B.Sc. degree in electrical engineering from the Technical College of Osnabrueck, Osnabrueck, Germany, in 1988, and his M.Sc. and Ph.D. degrees from Technical University of Berlin, Berlin, Germany, in 1992 and 1997, respectively. He is a Professor of the Department of Electrical Engineering and Head of the Electrical Networks Institute at Shahid Beheshti University. He was the Iran Research and Technology Institute for Electric Machines General Director for three years. His research interests include power system simulation, operation, planning, and control of power systems, renewable energy in power systems, and smart grids.