# Structure-Aware Recurrent Learning Machine for Short-Term Voltage Trajectory Sensitivity Prediction

Lipeng Zhu, *Member, IEEE,* Weijia Wen, Jiayong Li, *Member, IEEE,* Cong Zhang, *Member, IEEE,* Yangwu Shen, Yunhe Hou, *Senior Member, IEEE,* and Tao Liu, *Member, IEEE*

***Abstract*—In modern Internet of electric energy, i.e., networked power systems, data-driven schemes based on advanced machine learning methods have shown high potential in system emergency stability control, e.g., undervoltage load shedding (UVLS) against the short-term voltage stability (SVS) problem. However, how to efficiently and adaptively select the most effective UVLS sites for online SVS enhancement is still a challenging task. Faced with this issue, this paper develops an intelligent short-term voltage trajectory sensitivity index (VTSI) prediction scheme for adaptive UVLS site selection. Specifically, the scheme is realized by designing a powerful structure-aware recurrent learning machine (SRLM), which systematically combines the emerging graph convolutional network (GCN) with the recurrent long short-term memory algorithm. By doing so, the SRLM is not only fully aware of the non-Euclidean structure of the power grid, but also capable of amply capturing temporal features during SVS dynamics. Consequently, it manages to implement efficient and precise VTSI prediction, thereby reliably identifying critical UVLS sites in various scenarios. Numerical case studies on the Nordic test system illustrate the efficacy of the proposed scheme.**

***Index Terms*—Deep learning, emergency control, graph convolution, short-term voltage stability, trajectory sensitivity, undervoltage load shedding.**

## I. Introduction

As a representative paradigm of Internet of energy, modern electric power systems have been significantly upgraded with diverse advanced Internet of things (IoTs) technologies [1]–[3], yet it is still of paramount concern to ensure system security and stability during daily operation [4], [5]. Specifically, with the rapid growth of electric load demands in receiving-end areas, today's IoT-enabled power systems have to be tensely operated near their stability limits during peak-load hours. In urban load centers, with a rising proportion of dynamic loads, e.g., industrial motors and air conditioners, system operation is largely threatened by the short-term voltage stability (SVS) problem [6], [7]. Concretely, when transient faults occur, those dynamic loads with fast recovering behaviors would draw abnormally high reactive power from the system, leading to fault-induced delayed voltage recovery (FIDVR) [8], [9] or even voltage instability [10], [11]. Considering the dynamic loads' undesirable effect on inducing SVS issues, system-wide dynamic stability assessment needs to be reliably carried out soon after transient faults to predict whether potential instability would occur [12]–[14]. More importantly, remedial control actions should be rapidly taken once such emergency stability issues arise [15], [16].

In terms of emergency SVS control, among various kinds of corrective alternatives, undervoltage load shedding (UVLS) is the last resort but one of the most effective actions extensively used for system stabilization in practice [17]. The existing UVLS approaches for SVS control can be generally divided into the following two classes.

The first class is to shed loads based on fixed *if-then* rules [18]–[20]. For example, the threshold-based UVLS schemes widely utilized in practical grids curtail fixed proportions of loads by inspecting if individual bus voltages drop below a certain threshold over a preset time span [18]. Yet such schemes may be lack of adaptability and cause unnecessary load loss in practice due to fixed threshold settings in their UVLS rules. In recent years, some enhanced schemes have been developed to flexibly adjust the UVLS rules, e.g., dynamic UVLS amount assignment [19] and adaptive UVLS triggering via induction motor (IM) speed estimation [20]. However, as some of their parameters are still set to fixed values, they may have inadequate adaptability in complicated SVS scenarios. Recently, an adaptive UVLS scheme has been designed in [21] by designing particle swarm optimization (PSO)-based fuzzy controllers. while this scheme can be efficiently implemented, the PSO-based parameter setting may get trapped in local optima, leading to excessive load curtailment.

The other class belongs to data-driven solutions [9], [22]–[24]. In particular, with the availability of massive high-resolution system responsive data acquired by phasor measurement units (PMUs) [25], [26], advanced machine learning (ML) and/or deep learning (DL) technologies have been sufficiently leveraged to design data-driven UVLS schemes. To list a few, the oblique decision tree (DT) algorithm is employed in [22] to derive an adaptive UVLS scheme based on post-fault system responses. In [9], the kernel extreme learning machine

This work was supported in part by the National Natural Science Foundation of China under Grant 52207094, 52007056, 52377095, and 62088101, in part by the Natural Science Foundation of Hunan Province under Grant 2022JJ30007, in part by Hunan Key Laboratory for Internet of Things in Electricity under Grant 2019TP1016, in part by the Science and Technology Innovation Program of Hunan Province under Grant 2023RC3114, in part by the Research Grants Council of the Hong Kong Special Administrative Region under the Early Career Scheme through Project No. 27206021, and in part by the Guangdong Basic and Applied Basic Research Foundation under Grant 2023A1515011942. *(Corresponding Authors: Cong Zhang; Tao Liu.)*

Lipeng Zhu, Jiayong Li, and Cong Zhang are with the College of Electrical and Information Engineering, Hunan University, Changsha 410082, China (e-mail: zhulpwhu@126.com; j-y.li@connect.polyu.hk; zcong@hnu.edu.cn).

Weijia Wen is with the State Grid Hunan Information & Telecommunication Company, Changsha 410004, China (e-mail: cook_lex@foxmail.com).

Yangwu Shen is with the Wuhan Science and Technology Innovation Park of China Three Gorges Corporation, Wuhan 430021, China (e-mail: shenyangwu@126.com).

Yunhe Hou and Tao Liu are with the Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong, China (e-mail: yhhou@eee.hku.hk; taoliu@eee.hku.hk).

is exploited to carry out online UVLS on the basis of pre-fault system states. To implement confidence-aware UVLS, a DL-enabled Gaussian process learning model is proposed in [27]. Recently, by combining the DL technique with reinforcement learning, the deep reinforcement learning (DRL) method has been introduced to develop highly adaptive UVLS schemes [23], [24]. Compared with conventional rule-based schemes, these data-driven solutions have high potential to remarkably improve the reliability and adaptability of online UVLS in diverse emergency SVS situations.

Nonetheless, there exists a common defect in most of these UVLS schemes [18]–[20], [22]–[24], [27]. With no attention to the impact of shedding loads at different buses on system voltage recovery, they fix the sites (load buses) of UVLS execution. As a result, UVLS actions may be taken at buses with trivial effects on SVS enhancement, leading to a high risk of excessive load curtailment. Although the event-based UVLS scheme in [9] can implicitly determine the best UVLS sites for a presumed event, it might fail to do so when faced with new events. The fuzzy UVLS scheme in [21] can automatically determine the UVLS sites as well, yet it does not explicitly differ the criticality of different buses, which may trigger undue UVLS actions at too many load buses. To enhance the adaptability of online UVLS to diverse SVS events in practice, it is imperative to adaptively differ the criticality of individual load buses and thus select the most critical ones to help efficiently achieve voltage recovery with minimal load loss.

In fact, as verified in dynamic var planning against short-term voltage instability [28]–[30], the most effective control sites for SVS enhancement can be identified via voltage trajectory sensitivity (VTS) analysis. With load power at individual load buses perturbed one by one, the trajectory sensitivities of bus voltages to load change at each bus can be numerically estimated by observing the trajectory difference with/without perturbation. However, as multiple time-domain (T-D) simulations are needed to obtain the initial/perturbed voltage trajectories, the estimation procedure would induce high computational costs. Besides, with intricate structures and diverse unforeseen operating conditions in practical grids, the short-term VTS would change in different SVS scenarios. Both factors are likely to limit the applicability of numerical VTS analysis in emergency SVS control, where online decisions should be promptly and adaptively made in various complicated situations. Although VTS analysis can also be performed in an analytical way [31], the strict requirement on system model assumption could limit its capability and reliability in large-scale grids. In this regard, how to reliably and efficiently perform VTS estimation in complicated emergency situations remains an unsolved challenge in practice.

To tackle this challenging yet practical issue, this paper develops an intelligent short-term voltage trajectory sensitivity index (VTSI) prediction framework for adaptive UVLS implementation. Particularly, by combining the emerging graph convolutional network (GCN) with the promising long short-term memory (LSTM) algorithm, a powerful structure-aware recurrent learning machine (SRLM) is designed to predict VTSI in a highly efficient and reliable way. With both structural (spatial) characteristics and temporal features of the system fully learned by the SRLM, it is capable of correctly identifying the most influential load buses for UVLS against diverse SVS conditions and events. The main contributions and merits of this paper include:

1) This work for the first time develops an intelligent data-driven VTSI prediction framework that can be leveraged to efficiently and reliably select the most critical UVLS sites in diverse emergency situations. Hopefully, it would further boost the adaptability of existing UVLS schemes.
2) Being aware of the non-Euclidean structure of the given power grid, a novel SRLM is designed to fully learn the inherent spatial-temporal characteristics behind wide-area system dynamics. With such a powerful learning architecture, the SRLM is able to achieve highly reliable VTSI prediction in various complicated SVS scenarios.
3) As demonstrated by numerical experiments, the proposed framework has high potential in practical application contexts, being robust to unknown structural errors/changes and PMU measurement noises. Expectedly, these desirable features would make it more applicable in practice.

The remainder of the paper is organized as follows. The problem of VTS analysis and VTSI prediction for adaptive SVS control is first described in Section II. Then, Section III presents the details of the proposed SRLM. Afterward, extensive numerical tests on the Nordic test system are reported in Section IV to verify the SRLM's performances. Finally, concluding remarks are summarized in Section V.

## II. Problem Description

### A. Numerical VTS Analysis

Broadly speaking, trajectory sensitivity analysis is carried out by perturbing a specific dynamic system around its nominal trajectory [31]. With small changes (perturbations) imposed on system parameters, it is able to directly determine the trajectory deviation from the nominal one. This can further help characterize the relative importance of different parameters in changing system dynamics. Taking short-term VTS analysis for instance, it quantifies the sensitivities of voltage trajectories to active/reactive load power changes at individual buses [28]–[30]. By doing so, the most effective load buses can be selected as the control sites to mitigate the SVS problem in transient periods. In particular, given a receiving-end system with $M$ buses for SVS monitoring, the VTS of bus $i$ w.r.t. active power change at bus $j$ ($1 \leq i, j \leq M$) can be numerically estimated by

$$\frac{\partial V_i}{\partial p_j} \approx \frac{\boldsymbol{V}_i(p_{j0} + \Delta p_j, t) - \boldsymbol{V}_i(p_{j0}, t)}{\Delta p_j} \tag{1}$$

where $\Delta p_j$ is a small perturbation of active load power at bus $j$; $p_{j0}$ is the initial active load power at bus $j$; $\boldsymbol{V}_i(p_{j0}+\Delta p_j, t)$ and $\boldsymbol{V}_i(p_{j0}, t)$ are the voltage trajectory perturbed by $\Delta p_j$ and the initially nominal voltage trajectory with no perturbation. Note that, when the active load power at bus $j$ is perturbed, its counterpart, i.e., the reactive power at bus $j$, is perturbed proportionally. Unless otherwise stated, $\boldsymbol{V}_i(p_{j0} + \Delta p_j, t)$ and $\boldsymbol{V}_i(p_{j0}, t)$ will be briefly expressed as $\boldsymbol{V}_i^{\Delta p_j}(t)$ and $\boldsymbol{V}_i(t)$ in

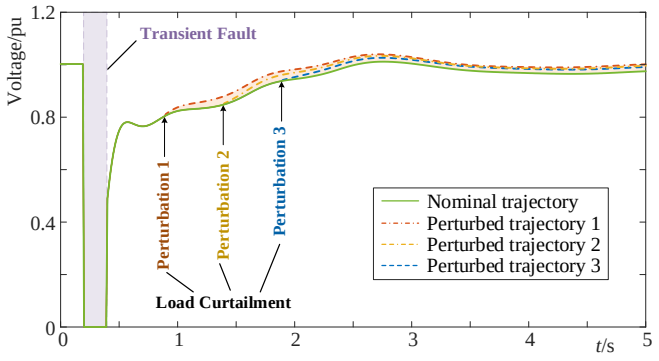



Fig. 1. Illustration of voltage trajectory sensitivity.

the sequel.

Here $\frac{\partial V_i}{\partial p_j}$ has a time-varying nature. If the time instant of imposing the perturbation varies, the perturbed trajectories will change accordingly. As shown in Fig. 1, with a small load curtailment of 50 MW at $t$ = 0.9 s, 1.4 s and 1.9 s, respectively, three distinct perturbed trajectories are observed, thus resulting in different VTS. In practical grids involving multiple rounds of load power changes, e.g., multi-round UVLS, the VTS needs to be re-calculated in new rounds of changes.

### B. VTSI Prediction for Adaptive SVS Control

Based on (1), the effect of the load change at bus $j$ on voltage recovery of bus $i$ in the transient time frame can be directly estimated. Further, given an observation time window (OTW) with length $T_{\text{win}}$ and a sampling interval $\Delta t$ for PMU-based time series (TS) trajectory acquisition, the VTSI w.r.t. $\Delta p_j$ can be calculated as [28], [30]:

$$VTSI_j \approx \frac{1}{MN_z}\sum_{z=1}^{N_z}\sum_{i=1}^{M}\left[\frac{\boldsymbol{V}_i^{\Delta p_j}(t)-\boldsymbol{V}_i(t)}{\Delta p_j}\right]_{t=t_z} \tag{2}$$

where $N_z = T_{\text{win}}/\Delta t$ is the total number of sampling points in the OTW. Intuitively, $VTSI_j$ can be deemed as the average area difference between the perturbed and nominal voltage trajectories (e.g., the shaded orange area in Fig. 1) of all the $M$ buses (in a normalized scale) within the OTW. In this regard, $VTSI_j$ can be used to estimate the overall impact of perturbing the load power at bus $j$ on the overall voltage recovery of the $M$ buses. If UVLS is taken as the major countermeasure for emergency SVS control, the VTSI values of all the buses can be calculated and sorted to select the most effective load buses for UVLS execution. However, if the VTSI is applied to online UVLS site selection in practical emergency situations, it would suffer from the following dilemma:

1) Due to the need for multiple T-D simulations to obtain perturbed/nominal voltage trajectories from each bus, the procedure of VTSI computation would incur heavy computational burdens. Consequently, this makes numerical VTSI calculation intractable in online contexts.
2) With the inevitability of online operation changes and unforeseen events, the VTSI values may vary from scenario to scenario. In this respect, online VTSI estimation rather than offline analysis should be carried out in real time, so as to adaptively respond to various unexpected SVS scenarios.

To address these tangled challenging issues, this paper develops an intelligent VTSI prediction scheme. Assisted by a well-designed SRLM, it is able to efficiently and reliably predict VTSI values in diverse SVS scenarios. How the SRLM is realized to implement VTSI prediction is detailed below.

## III. Proposed SRLM for VTSI Prediction

The main architecture of the proposed SRLM is illustrated in Fig. 2. It is carried out via two phases: 1) VTSI dataset preparation; and 2) structure-aware trajectory feature learning.

### A. VTSI Dataset Preparation

For the given receiving-end system with $M$ buses, considering various possible operating conditions and a representative contingency/event list provided by system operators, batch T-D simulations are performed to generate $n$ SVS scenarios. All the scenarios are integrated as an SVS case repository. Since these scenarios contain initial system trajectories with no perturbation, they are taken as nominal cases. For each nominal case, numerical perturbation is separately performed at each load bus to obtain a series of perturbed cases. Assuming there are $m$ ($m \le M$) load buses in the system, a small amount of load curtailment, denoted by $P_{\text{pert}}$, is taken at bus $j$ ($1 \le j \le m$), which results in $\Delta p_j = -P_{\text{pert}}$. Considering the potential need for multi-round UVLS-oriented SVS control in practice, the perturbation $\Delta p_j$ is set at different time instants, e.g., $T_1 = t_c + \Delta T$, $T_2 = t_c + 2\Delta T$ and $T_3 = t_c + 3\Delta T$ for three-round UVLS, where $t_c$ is the time instant of fault clearance and $\Delta T$ is the time interval between two successive UVLS actions. In this way, $3m$ perturbed cases are obtained from each nominal one via T-D simulations. All the perturbed cases are further gathered as a perturbed case repository.

After that, TS data acquisition and VTSI calculation are performed to collect a VTSI dataset from the above repositories. For a certain nominal case, three OTWs with length $T_{\text{win}} = \Delta T$ are employed to acquire multiple TS trajectories from the $M$ buses. Given the $k$th ($1 \le k \le 3$) OTW within the time slot of $T_k - \Delta T \sim T_k$, sequential responses of bus voltages, injected active power and reactive power, i.e., $\{V, P, Q\}$, are collected from each bus to form multiplex TS:

$$\begin{aligned}\boldsymbol{X}_k = &\{\boldsymbol{V}_1(t), \boldsymbol{V}_2(t), ..., \boldsymbol{V}_M(t), \boldsymbol{P}_1(t), \boldsymbol{P}_2(t), ..., \boldsymbol{P}_M(t),\\ &\boldsymbol{Q}_1(t), \boldsymbol{Q}_2(t), ..., \boldsymbol{Q}_M(t)\}\ (T_k - \Delta T \le t < T_k)\end{aligned} \tag{3}$$

where $\{\boldsymbol{V}_i(t), \boldsymbol{P}_i(t), \boldsymbol{Q}_i(t)\}$ are the nominal TS acquired from bus $i$ ($1 \le i \le M$). Then, voltage trajectories subject to the perturbation at $t = T_k$ are utilized to estimate the corresponding VTSI values. In particular, with the OTW set in the time span of $T_k \sim T_k + \Delta T$, VTSI values of the $m$ load buses are numerically calculated via (2) (with $T_k \le t_z < T_k + \Delta T$). For the remaining $(M - m)$ buses with no load, since no UVLS action can be taken as these buses, their VTSI values are set to 0 by default. All the buses' VTSI values estimated from the $k$th-round perturbation are collected as

$$\boldsymbol{Y}_k = [VTSI_{1k}, VTSI_{2k}, ..., VTSI_{Mk}] \tag{4}$$

Taking $\boldsymbol{X}_k$ and $\boldsymbol{Y}_k$ as the pairwise input and output for VTSI prediction, a VTSI instance is formed as $\{\boldsymbol{X}_k, \boldsymbol{Y}_k\}$. By processing all the $n$ SVS cases one by one, $3n$ VTSI

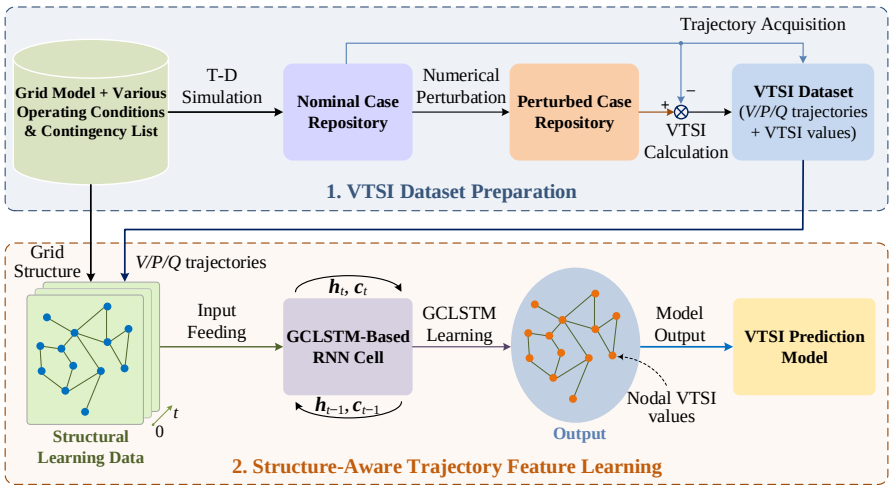



Fig. 2. Overall architecture of the proposed SRLM.

instances are produced and further gathered as a VTSI dataset. To eliminate the adverse effect of different ranges of individual variables on subsequent learning procedures, all the TS values of $\{V, P, Q\}$ in the inputs are scaled into the range of [0, 1] via min-max normalization. For VTSI values in the outputs, they are normalized into [-1, 1] in a different manner: all the positive and negative VTSI values are grouped together and separately mapped into the ranges of 0∼1 and -1∼0 based on min-max normalization. By doing so, the normalized VTSI would inherit the physical meaning of the raw values indicating the positive/negative effects of load changes on regional voltage recovery. The normalized dataset is represented as $\boldsymbol{\mathcal{S}} = \{(\boldsymbol{x}_i, \boldsymbol{y}_i) \mid \text{for } \boldsymbol{x}_i = \{\boldsymbol{x}_{it}\}_{t_1}^{t_{N_z}},\ 1 \le i \le 3n\}$.

### B. Structure-Aware Trajectory Feature Learning

Taking $\boldsymbol{\mathcal{S}}$ as the data source, in-depth feature learning is carried out to infer the mapping relationships between $\boldsymbol{x}_i$ and $\boldsymbol{y}_i$, based upon which a data-driven VTSI prediction model would be derived. In particular, by combining two powerful DL algorithms, i.e., GCN and LSTM, the inherent spatial-temporal features of the system are fully learned from regional TS trajectories in a structure-aware manner. Taking the voltage trajectories for instance, how spatial-temporal feature learning is implemented via GCN and LSTM is illustrated in Fig. 3. More details about how the two key DL algorithms GCN and LSTM are leveraged to realize spatial-temporal feature learning are presented below.

*1) GCN-Based Spatial Feature Learning:* Unlike the traditional convolutional neural network (CNN) specializing in learning spatial correlations in Euclidean domains, e.g., two-dimensional images, GCN is an emerging DL technique that has been demonstrated to be competent in handling irregular networked spaces or non-Euclidean graph structures [32]–[34]. In a practical power grid, as individual buses (nodes) are generally networked together via transmission lines in an irregular manner, it can be treated as a graph with complicated nodal dynamic behaviors and structural correlations between individual nodes [35]. From this viewpoint, the GCN algorithm naturally fits well with the need for grasping inherent structural characteristics of the power grid [36], [37]. Hence, it is utilized for structural feature learning in this paper.

First of all, by incorporating the data obtained in the 1st phase of VTSI dataset preparation with the network structure information, structural learning data to be fed into the GCN for spatial feature learning are formed. Specifically, for the $M$-bus power grid, its network structure is characterized as a graph with an adjacency matrix $\boldsymbol{A} = [a_{ij}]_{M\times M}$. In the given power grid, as the transmission line admittances represent the structural connections between individual buses [38]–[40], such admittance values are quite suitable for non-Euclidean network structure representation in GCN-based spatial feature learning. Thus, the entry $a_{ij}$ is estimated based upon admittance values:

$$a_{ij} = \begin{cases} |Y_{ij}|, & \text{if } i \neq j \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

Here $|Y_{ij}|$ is the magnitude of the mutual admittance of the transmission line connecting buses $i$ and $j$. Essentially, $\boldsymbol{A}$ can be taken as a bridge that connects the wide-area network structure with local dynamics at individual buses. As illustrated in Fig. 3, with the grid structure represented by an adjacency matrix-based graph, system-wide voltage values acquired from different buses' trajectories at three representative time instants ($t = 0$, 0.25 s, and 0.5 s, respectively) can be correspondingly represented by three graphs. With different nodal colors standing for different voltage values at respective buses, the three graphs present structural voltage distributions of the grid. By feeding such graph data obtained from different time instants into a GCN, system-wide spatial features w.r.t. networked voltage distributions would be learned. Before detailing how the GCN is constructed, the concept of graph Laplacian that enables spectral graph convolution is first introduced here. In particular, based on the adjacency matrix $\boldsymbol{A}$, the normalized graph Laplacian is described by

$$\boldsymbol{L} = \boldsymbol{I}_M - \boldsymbol{D}^{-1/2}\boldsymbol{A}\boldsymbol{D}^{-1/2} \tag{6}$$

where $\boldsymbol{I}_M$ is the identity matrix of size $M$; $\boldsymbol{D}$ is a diagonal matrix with $D_{ii} = \sum_j a_{ij}$. Let the eigenvalues of

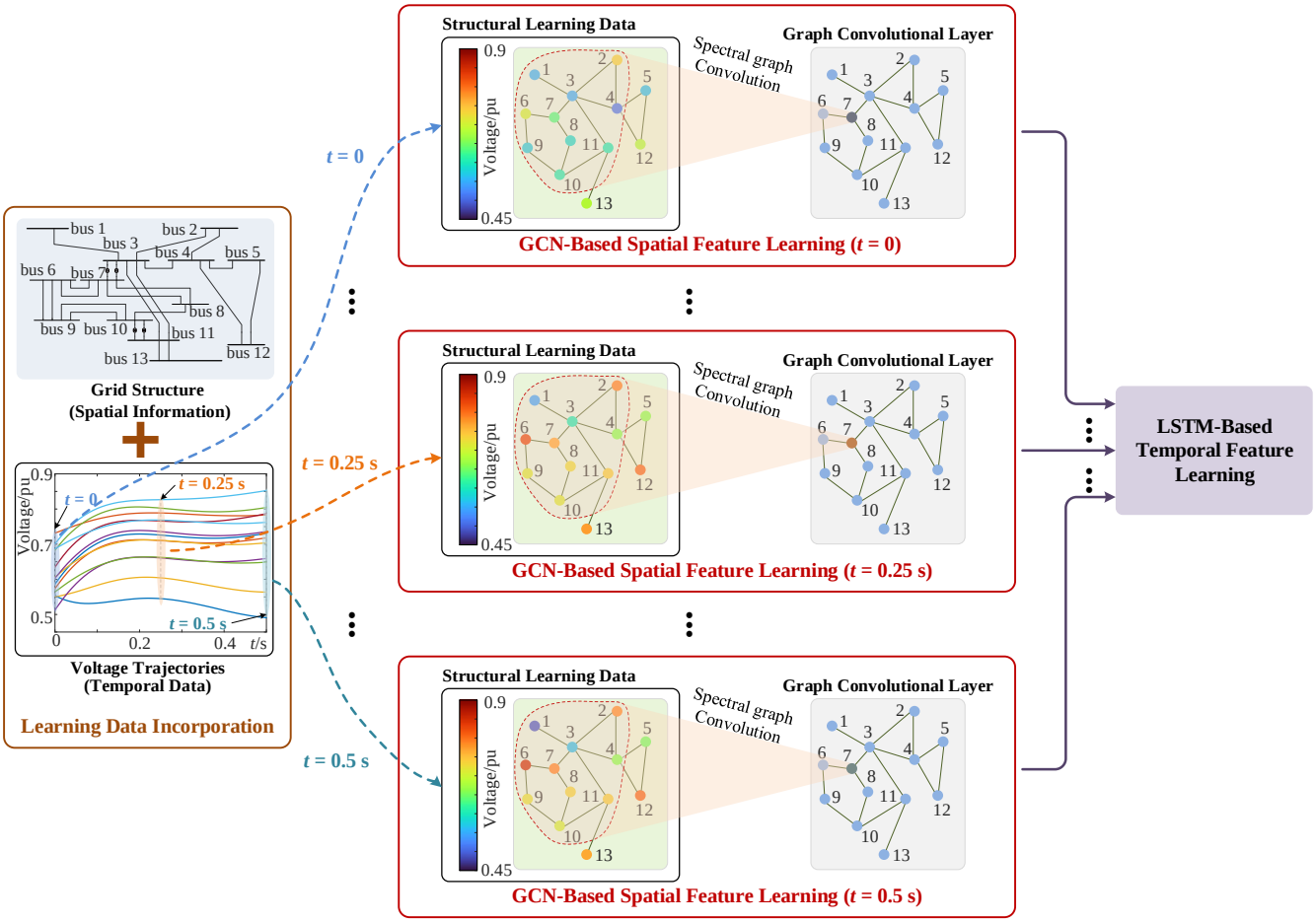



Fig. 3. Illustration of spatial-temporal feature learning with GCN and LSTM.

$\boldsymbol{L}$ be denoted as $\{\lambda_l\}_0^{M-1}$, and the corresponding eigenvectors are represented by $\{\boldsymbol{u}_l\}_0^{M-1}$. From the perspective of graph signal processing [41], the matrix of eigenvectors $\boldsymbol{U} = [\boldsymbol{u}_0, \boldsymbol{u}_1, ..., \boldsymbol{u}_{M-1}]$ can be deemed as the Fourier basis to diagonalize $\boldsymbol{L}$, which yields $\boldsymbol{L} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^T$, where $\boldsymbol{\Lambda} = \text{diag}([\lambda_0, \lambda_1, ..., \lambda_{M-1}])$. For a certain snapshot of the $M$-bus $\{V, P, Q\}$ TS responses, which is denoted as $\boldsymbol{v} \in \mathbb{R}^{M\times 3}$, $\hat{\boldsymbol{v}} = \boldsymbol{U}^T\boldsymbol{v}$ and $\boldsymbol{v} = \boldsymbol{U}\hat{\boldsymbol{v}}$ are taken as the graph Fourier transform and the inverse counterpart [41] on the grid. Following this concept, spectral graph convolution on the graph signal $\boldsymbol{v}$ is conducted via spectral filtering [32]:

$$\boldsymbol{\mathcal{V}} = g_\theta \star \boldsymbol{v} = g_\theta(\boldsymbol{L})\boldsymbol{v} = g_\theta(\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^T)\boldsymbol{v} = \boldsymbol{U}g_\theta(\boldsymbol{\Lambda})\boldsymbol{U}^T\boldsymbol{v} \quad (7)$$

where $\boldsymbol{\mathcal{V}} \in \mathbb{R}^{M\times 3}$ is the convolution result; $g_\theta = \text{diag}(\theta)$ is a graph filter, with $\theta$ denoting a group of Fourier coefficients. Essentially, $g_\theta$ can be understood as a function of the eigenvalues of $\boldsymbol{L}$, i.e., $g_\theta(\boldsymbol{\Lambda})$. Considering the heavy computational burden of the raw graph convolution in (7), a simplified version with much higher computational efficiency is proposed in [32] by approximating $g_\theta(\boldsymbol{L})$ with truncated Chebyshev polynomials:

$$\boldsymbol{\mathcal{V}} = g_\theta \star \boldsymbol{v} = g_\theta(\boldsymbol{L})\boldsymbol{v} \approx \sum_{k=0}^{K-1} \theta_k T_k(\tilde{\boldsymbol{L}})\boldsymbol{v} \quad (8)$$

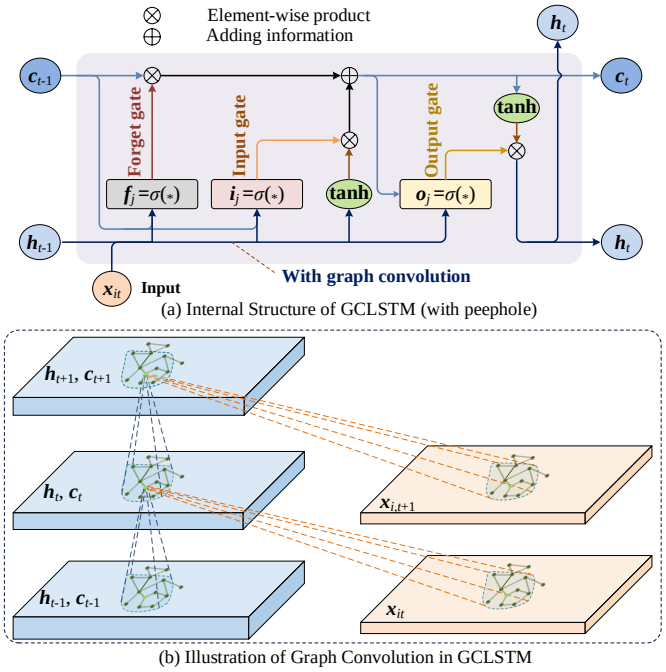
where $\theta_k \in \mathbb{R}$ and $T_k(\tilde{\boldsymbol{L}}) \in \mathbb{R}^{M\times M}$ are the $k$th-order Chebyshev coefficient and polynomial, respectively; $\tilde{\boldsymbol{L}} = 2\boldsymbol{L}/\lambda_{\max} - \boldsymbol{I}_M$ is the normalized Laplacian. The Chebyshev polynomial $T_k(\tilde{\boldsymbol{L}})$ can be recursively computed by

$$T_k(\tilde{\boldsymbol{L}}) = \begin{cases} 1, & k = 0 \\ \tilde{\boldsymbol{L}}, & k = 1 \\ 2\tilde{\boldsymbol{L}}T_{k-1}(\tilde{\boldsymbol{L}}) - T_{k-2}(\tilde{\boldsymbol{L}}), & k \geq 2 \end{cases} \quad (9)$$

Note that the approximate graph convolution in (8) has a $K$-localized receptive field, where $K$ denotes the kernel size. Taking $K = 3$ for instance, as shown in Fig. 3, the information collected from the buses with at most $K - 1 = 2$ hops from the central bus is utilized for graph convolution. Based on (8), a GCN can be constructed by connecting multiple graph convolution neurons together to sufficiently learn the inherent structural correlations within the grid.

After GCN-based spatial feature learning from individual graphs, the learned hidden states at different time instants are fed into LSTM-based neurons to further mine temporal features hidden behind system-wide sequential trajectories, as presented in what follows.

*2) LSTM-Based Temporal Feature Learning:* The above graph convolution focuses on structural feature learning from single snapshots of the grid's dynamics. To capture system temporal characteristics behind TS trajectories, the LSTM algorithm with high potential in temporal representation learning




(a) Internal Structure of GCLSTM (with peephole)

(b) Illustration of Graph Convolution in GCLSTM

Fig. 4. Structure of the composite GCLSTM.

from TS data [42] is further considered for feature learning. Specifically, by incorporating the GCN technique with the LSTM algorithm [34], a composite method called graph convolutional LSTM (GCLSTM) is introduced in this paper. As shown in Fig. 4(a), a GCLSTM cell consists of an input gate $\boldsymbol{i}_t$, an output gate $\boldsymbol{o}_t$, a memory cell $\boldsymbol{c}_t$ and a forget gate $\boldsymbol{f}_t$. All of these components collaborate with each other to derive sequential hidden state $\boldsymbol{h}_t$ from the temporal input $\boldsymbol{x}_{it}$. In fact, it shares similar internal working mechanisms with a standard LSTM unit, except that graph convolution is operated on $\boldsymbol{x}_{it}$ and $\boldsymbol{h}_t$ to yield structure-aware $\boldsymbol{h}_{t+1}$ and $\boldsymbol{c}_{t+1}$, as depicted in Fig. 4(b). Mathematically, the comprehensive operation in the GCLSTM cell can be described as

$$\boldsymbol{i}_t = \sigma\left(\boldsymbol{W}_{xi} \star \boldsymbol{x}_{it} + \boldsymbol{W}_{hi} \star \boldsymbol{h}_{t-1} + \boldsymbol{W}_{ci}\boldsymbol{c}_{t-1} + \boldsymbol{b}_i\right) \tag{10}$$

$$\boldsymbol{f}_t = \sigma\left(\boldsymbol{W}_{xf} \star \boldsymbol{x}_{it} + \boldsymbol{W}_{hf} \star \boldsymbol{h}_{t-1} + \boldsymbol{W}_{cf}\boldsymbol{c}_{t-1} + \boldsymbol{b}_f\right) \tag{11}$$

$$\boldsymbol{c}_t = \boldsymbol{f}_t \odot \boldsymbol{c}_{t-1} + \boldsymbol{i}_t \odot \tanh\left(\boldsymbol{W}_{xc} \star \boldsymbol{x}_{it} + \boldsymbol{W}_{hc} \star \boldsymbol{h}_{t-1} + \boldsymbol{b}_c\right) \tag{12}$$

$$\boldsymbol{o}_t = \sigma\left(\boldsymbol{W}_{xo} \star \boldsymbol{x}_{it} + \boldsymbol{W}_{ho} \star \boldsymbol{h}_{t-1} + \boldsymbol{W}_{co}\boldsymbol{c}_t + \boldsymbol{b}_o\right) \tag{13}$$

$$\boldsymbol{h}_t = \boldsymbol{o}_t \odot \tanh\left(\boldsymbol{c}_t\right) \tag{14}$$

where $\boldsymbol{W}_{x*}$ and $\boldsymbol{W}_{h*}$ $(* \in \{i, f, c, o\})$ are the Chebyshev coefficient matrices to be learned; $\star$ denotes the graph convolution described in (8); $\boldsymbol{W}_{c*}$ $(* \in \{i, f, o\})$ is the weight matrix linking $\boldsymbol{c}_t$ to other components ; $\boldsymbol{b}_*$ $(* \in \{i, f, c, o\})$ stands for the bias parameters; $\odot$ represents element-wise product; $\sigma(\kappa) = 1/(1 + e^{-\kappa})$ is the sigmoid activation function.

In fact, the GCLSTM can be regarded as a variation of the base LSTM, where the conventional multiplication by weight matrices is replaced by graph convolution with kernels in the form of Chebyshev coefficient matrices. In this regard, graph convolution is essentially embedded into the LSTM unit to perform convolution for the structural learning data and the hidden states at different time instants during LSTM-enabled temporal correlation learning. With such a combinatorial learning mechanism, the GCLSTM not only learns temporal correlations from the sequential inputs, but also makes the learning procedure be aware of the structural (spatial) characteristics of the grid. As a result, the hidden state $\boldsymbol{h}_t$ would comprehensively capture networked spatial-temporal features from the structural system dynamics.

To enhance the learning capability, multiple GCLSTM layers can be stacked for in-depth feature learning. A fully connected layer is further attached behind the GCLSTM layers to map the hidden states to the eventual output, i.e., VTSI values. Let the mapping relationship along all these neural layers be represented by $\mathcal{H}(\boldsymbol{x}_i)$. A linear activation function is employed to derive the ultimate prediction, i.e., $\hat{\boldsymbol{y}}_i = \mathcal{H}(\boldsymbol{x}_i)$. With emphasis on regressive prediction, the objective of the whole network is to minimize the per-instance loss function below:

$$\min \mathcal{L} = \min \left(\hat{\boldsymbol{y}}_i - \boldsymbol{y}_i\right)^2 \tag{15}$$

To augment the whole SRLM's generalization performance, the dropout trick randomly dropping out neural units [42] is introduced into the GCLSTM training procedure. After the training procedure is finished, the SRLM derives a data-driven VTSI prediction model for online VTSI prediction.

### C. Online VTSI Prediction

During online SVS monitoring, an OTW with both its length and sliding step set to $\Delta T$ is employed to sequentially acquire TS trajectories from individual buses via PMUs after fault clearance. Assuming a multi-round UVLS scheme is readily available for SVS control, the SRLM is dedicated to assisting the scheme via adaptive UVLS site selection. Specifically, with multiple windows of $\{V, P, Q\}$ TS separately collected and fed into the SRLM, it predicts VTSI values before each round of UVLS execution. For a certain UVLS round, all the load buses with positive VTSI prediction values are sorted in descending order, and the first $m'$ $(m' < m)$ buses are selected as the control sites for load curtailment. As system responses in different OTWs may lead to different VTSI values, the SRLM would be able to enhance the UVLS scheme's performance by adaptively selecting the most critical control sites.

For the sake of evaluating the prediction performances of the SRLM, three statistical metrics are calculated. In particular, with $n_{\text{test}}$ SVS cases provided for online test, three-round UVLS actions result in $N = 3n_{\text{test}}$ VTSI instances in total. Let the predicted and actual VTSI values of instance $i$ $(1 \le i \le N)$ be $\hat{\boldsymbol{y}}_i = [\hat{y}_{i1}, \hat{y}_{i2}, ..., \hat{y}_{iM}]$ and $\boldsymbol{y}_i = [y_{i1}, y_{i2}, ..., y_{iM}]$, the mean absolute error (MAE), root mean square error (RMSE) and reliability of UVLS site selection are estimated as

$$MAE = \frac{1}{MN}\sum_{i=1}^{N}\sum_{j=1}^{M}\left|\hat{y}_{ij} - y_{ij}\right| \tag{16}$$

$$RMSE = \sqrt{\frac{1}{MN}\sum_{i=1}^{N}\left(\hat{\boldsymbol{y}}_i - \boldsymbol{y}_i\right)^2} \tag{17}$$

$$R_{el} = \frac{1}{N}\sum_{i=1}^{N}\psi\left[\phi(\hat{\boldsymbol{y}}_i, m') \cap \phi(\boldsymbol{y}_i, m')\right]/m' \times 100\% \tag{18}$$

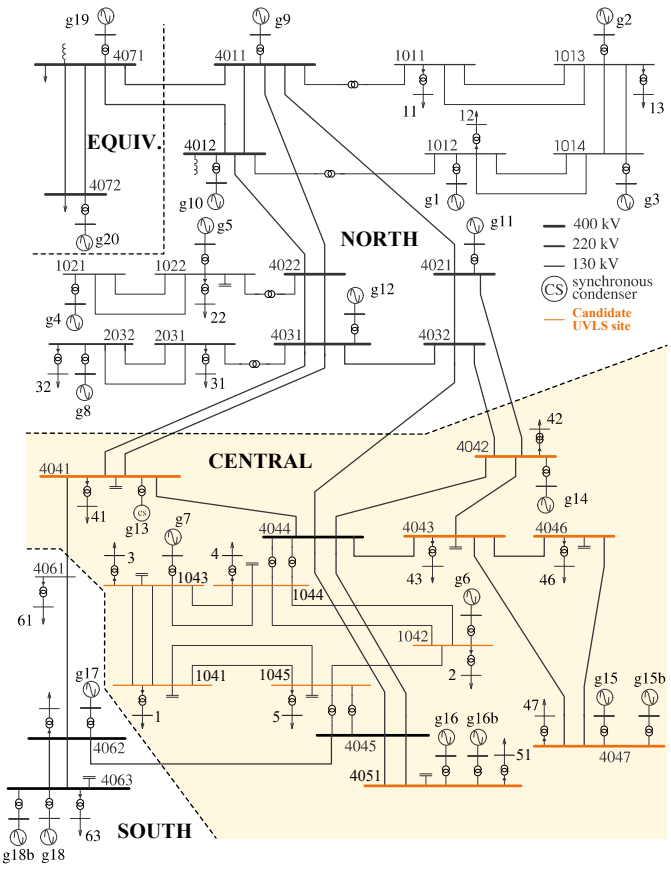



Fig. 5. Single-line diagram of the Nordic test system.

where $\phi(*, m')$ is the sorting function retrieving the indices of the $m'$ load buses with the largest positive VTSI values, and $\psi[\boldsymbol{a}]$ is a counting function that counts the total number of elements in the set $\boldsymbol{a}$.

## IV. Case Study

### A. Simulation Setting and Case Generation

The comprehensive performance of the SRLM-based prediction scheme was tested on the Nordic test system. This is a simplified 77-bus system of the actual Swedish and Nordic power grid [43], as shown in Fig. 5. It has been recently recommended by an IEEE PES task force for voltage stability research [44]. Following the system description in [43], the system was modeled and numerically simulated in PSD-BPA, a commercial power grid simulation package extensively used in China. Focusing on the SVS issue, all the loads in the system were modeled as composite loads with IM and static ZIP loads in parallel. In the system, the heavily loaded area called Central (see the shaded area in Fig. 5) consumes more than 55% of the total electric power, and it was chosen as the major receiving-end region for SVS monitoring. Assuming all the 13 130-kV and 400-kV buses in this area were deployed with PMUs, TS responses of $\{V, P, Q\}$ in SVS dynamics were sequentially acquired from these buses via PMUs during online monitoring. Among them, the 11 load buses were taken as candidate UVLS sites, as highlighted in Fig. 5.

Taking into account various representative operating points (50 ones), typical dynamic load proportions (60%, 70%, 80%), diverse transient fault locations (15 ones) and different fault clearing times ($t_c$ = 0.2 s or 0.3 s), 4500 SVS scenarios were generated via T-D simulations in PSD-BPA. For each scenario, three shots of post-fault perturbations in the form of load curtailment were separately applied to the 11 load buses for VTSI estimation. With pairwise $\{V, P, Q\}$ trajectories and VTSI values extracted from all the scenarios, 13500 VTSI instances were obtained. The main parameters during case generation were set to: $M$ = 13, $m$ = 11, $P_{\text{pert}}$ = 50 MW, $\Delta t$ = 0.01 s, $T_{\text{win}} = \Delta T$ = 0.5 s and $N_z$ = 50.

TABLE I
Primary Parameter Setting for SRLM Training

| Parameter | Symbol | Value |
|---|---|---|
| Kernel size for graph convolution | $K$ | 3 |
| No. of nodes for graph convolution | $M$ | 13 |
| No. of neurons in the 1st GCLSTM layer | $n_{L1}$ | 64 |
| No. of neurons in the 2nd GCLSTM layer | $n_{L2}$ | 32 |
| Learning rate | $\alpha$ | 0.001 |
| Dropout probability | $p_{drop}$ | 50% |
| Batch size | $n_B$ | 64 |
| No. of epochs | $n_E$ | 500 |

The above VTSI dataset with 13500 cases was divided into three groups, including a training set for offline learning (9000 cases), a validation set for offline validation (2000 cases) and a testing set for online testing (2500 cases). Note that the testing set was separated from the whole dataset by deterministic splitting. For the remaining cases, they were partitioned into the training and validation sets via random sampling with no replacement. In this way, the testing cases would contain SVS scenarios that are totally unseen by offline learning, so as to better demonstrate the generalization performances of the SRLM in unforeseen conditions. For SRLM construction, two GCLSTM layers were stacked to build the learning machine for VTSI prediction. In terms of SRLM training, the learning parameters were set on the basis of the authors' empirical tests, with the primary parameter setting summarized in Table I. The RMSProp algorithm [42] was employed as the optimizer to implement iterative SRLM training.

### B. Comprehensive Learning Performances

*1) Illustration of VTSI Prediction:* Before verifying the overall performance of the SRLM, how it performs on VTSI prediction is illustrated here with a case randomly chosen from the testing set. The actual and predicted VTSI values of this case are presented in Fig. 6. As can be seen, the SRLM achieves highly precise VTSI prediction, with all the predicted values slightly deviating from the actual ones. In addition, the relative ranking order of all the predicted VTSI values remains 100% consistent with the actual order. Hence, when UVLS sites are determined on the basis of the sorting order of predicted VTSI values, those buses with high positive impacts on regional voltage recovery would be reliably selected.

*2) Overall Prediction Results:* With each of the training/validation/testing cases fed into the SRLM, its prediction performances are comprehensively evaluated, as summarized in Table II. For the reliability metric, the number of buses

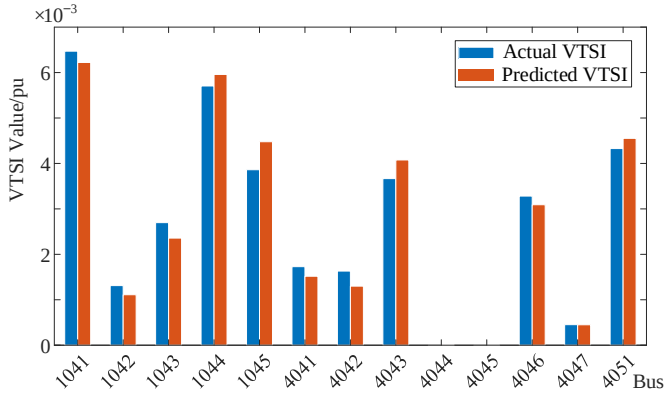



Fig. 6. VTSI prediction result of a randomly chosen case (with VTSI values rescaled into the raw range).

TABLE II
COMPREHENSIVE VTSI PREDICTION PERFORMANCES

| Dataset | $MAE$ | $RMSE$ | $R_{el}/\%$ |
|---|---|---|---|
| Training set (9000) | 0.0114 | 0.0168 | 98.58 |
| Validation set (2000) | 0.0117 | 0.0170 | 98.21 |
| Testing set (2500) | 0.0118 | 0.0173 | 97.96 |

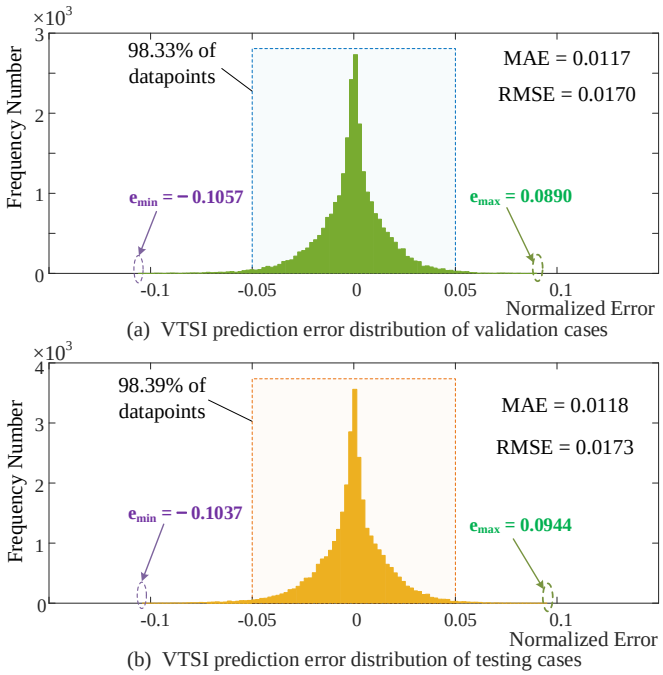



Fig. 7. Error distribution of VTSI prediction.

selected for UVLS was set to $m' = 5$. Clearly, the SRLM is able to maintain high prediction precision on all the three datasets, with the MAE and RMSE indices kept below 0.012 and 0.018, respectively. As for assisting UVLS site selection, the critical load buses can be correctly selected with a high reliability of around 98% even for the testing cases not seen by offline learning. Further, the prediction error distributions of all the validation and testing cases are statistically drawn in Fig. 7. Over 98% of the prediction errors fall into the range of [-0.05, 0.05], with a highly concentrated distribution around 0. For the worst cases, the largest absolute error is less than 11% of the normalized VTSI range (0∼1). These distributions again imply the SRLM's desirable VTSI prediction performances.

*3) Comparative Study:* For the purpose of further illustrating the advantage of the proposed SRLM, a comparative study was carried out here by comparing it with some representative DL/ML methods. In particular, three conventional ML methods, including multilayer perceptron (MLP) based neural network, support vector machine (SVM) and classification and regression tree (CART), and three representative DL algorithms, i.e., CNN, LSTM, and ConvLSTM (with a combination of conventional CNN and LSTM) [45], were also implemented for VTSI prediction. Since the first three comparative methods with shallow learning structures cannot handle multi-dimensional TS data, all the sampling values in a given OTW were taken as independent dimensions of inputs for regressive learning. Besides, as their outputs are limited to single values, specific VTSI prediction models have to be separately customized for each of the load buses. For the CNN, LSTM, and ConvLSTM algorithms, as they are compatible with multi-dimensional TS inputs and vectorized outputs, no additional data preprocessing was needed. All the training/validation cases generated in Section IV-A were fed into these methods for offline training, and their online prediction performances on the same testing set with 2500 unforeseen cases are summarized in Table III.

TABLE III
COMPARISON OF VTSI PREDICTION PERFORMANCES ON TESTING CASES

| Method | $MAE$ | $RMSE$ | $R_{el}/\%$ |
|---|---|---|---|
| Proposed | 0.0118 | 0.0173 | 97.96 |
| MLP | 0.0278 | 0.0390 | 87.71 |
| SVM | 0.0269 | 0.0386 | 89.19 |
| CART | 0.0296 | 0.0424 | 83.87 |
| CNN | 0.0210 | 0.0323 | 92.50 |
| LSTM | 0.0203 | 0.0311 | 94.17 |
| ConvLSTM | 0.0189 | 0.0292 | 96.64 |

Evidently, none of these competitors defeat the proposed SRLM. Specifically, the VTSI prediction errors of the three shallow learners, i.e., MLP, SVM and CART, are more than 120% larger than that of the SRLM, leading to a low reliability of less than 90%. This is mainly caused by their learning manner that simply treats all the sampling values as independent features. Consequently, the inherent spatial-temporal correlations between networked buses during SVS dynamics are totally omitted, thus resulting in degraded VTSI prediction performances. On the other hand, the CNN, LSTM, and ConvLSTM methods perform better than those shallow learners, with the overall reliability remaining above 92%. Among these three DL-based alternatives, ConvLSTM achieves the best VTSI prediction performance, being able to reliably identify more than 96.5% of the critical UVLS sites. In fact, this is because its deep neural structures with the combination of CNN and LSTM are able to learn the spatial-temporal correlations between successive sampling points of wide-area dynamics in the specific power grid. Nonetheless, as such correlations do not naturally capture the inherent spatial-

TABLE IV
STATISTICS OF COMPUTATIONAL EFFICIENCIES OF DIFFERENT METHODS

| Method | Offline training time/s | Offline computational memory/MB | Model storage space/MB | Online VTSI estimation time/s |
|---|---|---|---|---|
| Proposed | 9897.5237 | 2543.793 | 16.887 | 0.0038 |
| MLP | 69.5852 | 96.862 | 2.458 | 0.0021 |
| SVM | 61.6371 | 89.389 | 3.646 | 0.0023 |
| CART | 48.5407 | 73.735 | 2.682 | 0.0022 |
| CNN | 1571.5113 | 1183.043 | 5.168 | 0.0030 |
| LSTM | 2868.3285 | 1363.842 | 7.872 | 0.0033 |
| ConvLSTM | 4332.3744 | 2086.961 | 13.284 | 0.0034 |
| Numerical | – | – | – | 102.2883 |

***Note 1***: The 2nd column refers to the offline time consumption of data-driven VTSI prediction model training; the 3rd column shows the computational memory usage during VTSI prediction model training; the 4th column represents the consumed storage space to save the trained VTSI prediction model; the 5th column denotes the average computation time of online VTSI estimation for the 2500 testing cases.

temporal characteristics within the non-Euclidean structural grid, it still falls behind the SRLM by a certain degree. This also reveals the potential of graph learning in enhancing structural feature learning in practical power grids. Since practical grids often present complicated spatial-temporal dynamics, the SRLM sufficiently learning both structural characteristics and temporal correlations would be more applicable in practice.

*4) Computational Efficiency:* To evaluate the computational efficiency of the SRLM, its time and space complexities during both offline model training and online application for VTSI prediction were tested with four metrics, including offline training time, offline computational memory, model storage space, and online VTSI estimation time (see *Note* 1 below Table IV for detailed explanations) [46]. In addition to the SRLM, the computational efficiencies of the six comparative methods involved in Table III were also examined here for comparison. All the computation tasks were conducted in Python 3.6 with the TensorFlow backend. A PC configured with a 3.60-GHz∗8 Intel Core i7-7700 CPU and an NVIDIA GeForce GTX-1080 GPU was used to fulfill these computation tasks. To verify the advantage of these ML-based methods in enhancing the online application efficiency, T-D simulation-based numerical VTS estimation that directly estimates VTSI values for online testing cases in a numerical way was also implemented. Table IV summarizes the computational efficiency test results of all the methods.

Overall, the proposed SRLM has the highest offline time and space complexities among all the ML-based methods. Yet it should be mentioned that the SRLM's offline computational complexities are still affordable, with the training procedure completed in no more than 2.8 hours and the computational memory consumption being less than 2.6 GB. Besides, despite the fact that the offline training efficiency of the SRLM is not very high, this does not essentially affect its online application efficiency. In fact, with its prediction model occupying less than 20 MB for storage, it can perform online VTSI prediction in a highly efficient way by directly feeding TS data into the prediction model. As a result, it issues the VTSI prediction

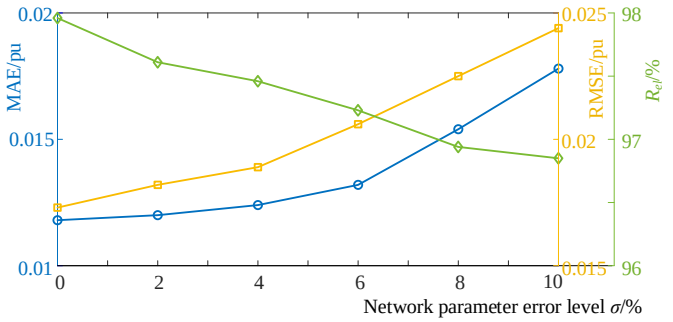



Fig. 8. VTSI prediction performance with network parameter errors.

results in less than 4 ms on average after the completion of TS data acquisition. This would largely contribute to timely UVLS decision making in urgent conditions. In contrast, although the numerical VTS analysis alternative does not involve offline computation, it takes nearly 2 min to estimate the VTSI values for a single online testing case via multiple T-D simulations. For the remaining ML-based schemes, they have highly competitive online application efficiencies and relatively low offline computational complexities. Yet considering that the proposed SRLM is able to achieve more reliable VTSI prediction without essentially sacrificing the online efficiency, the SRLM would be more helpful in practical emergency situations.

### C. Impact of Structural Discrepancies

As power network modeling errors are often inevitable, the network discrepancies between offline simulation environments and online operation contexts in practice cannot be ignored. Taking into account such structural discrepancies, random errors following the normal distribution $N(0, \sigma_1^2)$ were imposed onto transmission line parameters during offline case generation. In particular, the standard deviation $\sigma_1$ was set to $\sigma_1 = 2\%$, $4\%$, $6\%$, $8\%$ and $10\%$ to simulate different levels of line parameter errors. The adjacency matrix $\boldsymbol{A}$ for graph convolution [see Section III-B.1)] was re-calculated accordingly to account for such errors. The SRLM was re-trained with cases generated from such discrepant networks, and its VTSI prediction performances in practical online contexts represented by the testing cases with no structural errors are depicted in Fig. 8. As can be observed, the SRLM manages to maintain satisfactory performances against structural errors. Even if the error level is as high as 10%, it is able to select the most critical UVLS sites with high reliability ($> 96.5\%$).

Moreover, considering unforeseen structural discrepancies caused by network changes during online monitoring, the SRLM's online performances were further tested with unknown topological changes. In particular, two groups of $N-1$ and $N-2$ topological changes were generated by randomly disconnecting transmission lines in the system. Following the simulation settings in Section IV-A, 2500 new testing cases were produced for each of the groups. The SRLM's VTSI prediction performances on these new testing cases are summarized in Table V. It is found that the SRLM remains relatively robust to such topological changes, with the overall

TABLE V
VTSI PREDICTION PERFORMANCES WITH TOPOLOGICAL CHANGES

| Condition | $MAE$ | $RMSE$ | $R_{el}/\%$ |
|---|---|---|---|
| Normal | 0.0118 | 0.0173 | 97.96 |
| $N-1$ | 0.0123 | 0.0181 | 97.43 |
| $N-2$ | 0.0188 | 0.0306 | 94.86 |

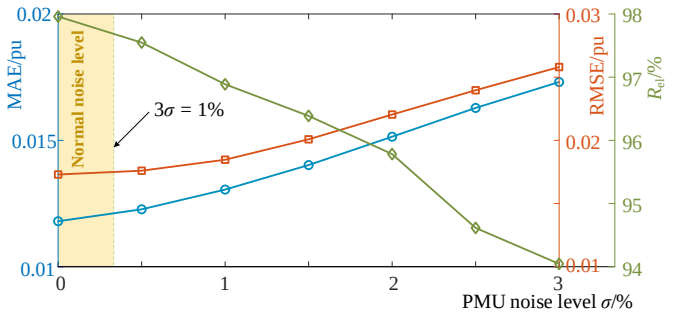



Fig. 9. VTSI prediction performance with PMU measurement noises.

reliability kept above 97% and 94.5% in totally unknown $N-1$ and $N-2$ conditions. This verifies that the SRLM can tolerate relatively severe structural discrepancies during online application, being competent in practical changing conditions.

### D. Robustness to PMU Measurement Noises

Considering the unavoidable PMU measurement noises in practice, the performances of the SRLM under practical noisy measurement contexts were examined here. Specifically, PMU measurement noises were generated by adding random noises with distribution $N(0,\sigma_2^2)$ to the raw $\{V,P,Q\}$ TS of the 2500 online testing cases produced in Section IV-A. By specifying $\sigma_2$ as 0.5%, 1%, 1.5%, 2%, 2.5% and 3%, different levels of measurement noises were simulated. Fig. 9 presents the SRLM's performances under these noise levels. As can be found, the SRLM can hold acceptable VTSI prediction precision under different noise levels. Even for the extremely stringent condition with $\sigma_2 = 3\%$ (corresponding to absolute noises as large as $\pm 9\%$ based on the $3\sigma$ rule), it can maintain the reliability above 94%. In fact, if the PMU measurement precision is in accordance with the normal IEEE standard [47], the SRLM would be able to firmly keep its reliability above 97.5%. Hence, the SRLM is expected to work well in practical measurement contexts with normal noise levels.

### E. Potential in Enhancing Online SVS Control

The above tests have exhibited the excellent performances of the proposed SRLM on online VTSI prediction. To further demonstrate its potential in helping enhance online SVS control, VTSI-assisted UVLS is illustrated with a randomly chosen online testing case. For simplicity, the threshold-based UVLS scheme shedding fixed proportions of loads (8%, 12% and 12% for three rounds) [20] was adopted here. Note that, instead of curtailing loads at all the load buses, which often causes excessive load loss in practice, here UVLS is executed at the $m'$ ($m'$ = 5) load buses with the largest positive VTSI prediction values. Meanwhile, the UVLS strategy that selects

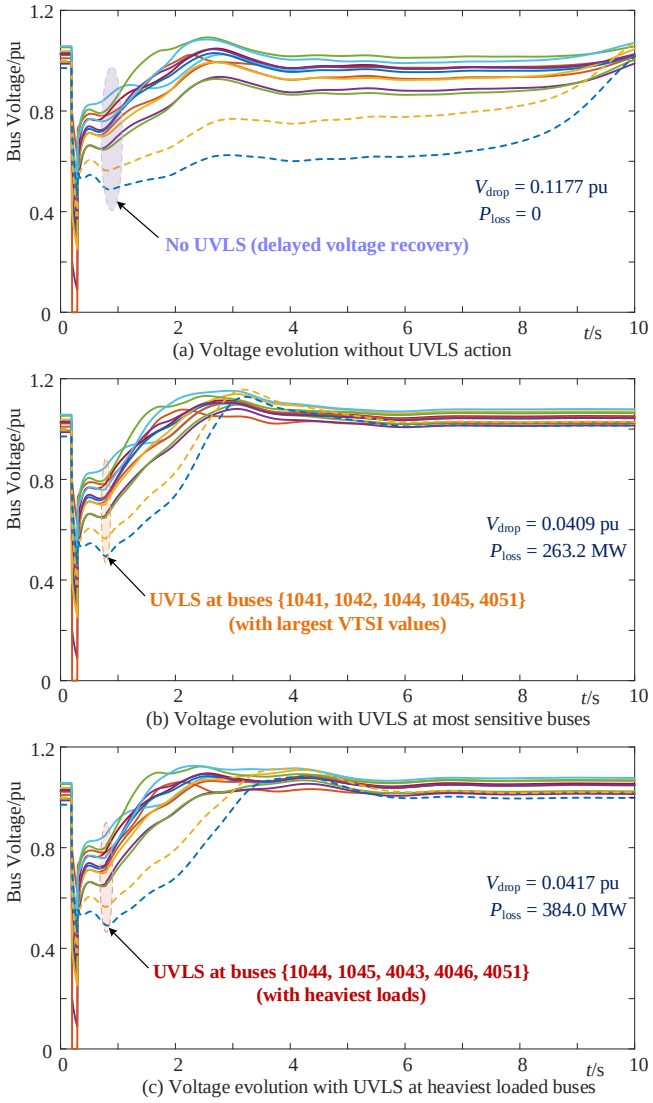



Fig. 10. Representative example of UVLS with the help of VTSI-based control site selection. ($V_{drop}$ → average voltage drop w.r.t. steady-state voltage levels at individual buses; $P_{loss}$ → total amount of load curtailment)

the five buses with the heaviest loads to perform UVLS was considered for comparison. Fig. 10 depicts the voltage profiles of the chosen testing case with different UVLS strategies.

Without UVLS action, the system experiences FIDVR after transient fault clearance, with an average voltage drop of more than 0.1 pu. With one-round UVLS actions, all the buses' voltages are recovered at a much higher speed, as shown in Fig. 10(b)-(c). Yet it should be noticed that the UVLS strategy based on VTSI prediction is able to better mitigate transient voltage drops while curtailing 32% fewer loads than the strategy shedding loads at heaviest loaded buses. With special attention to the two buses undergoing the lowest voltage levels (see dotted lines in Fig. 10), it can be seen that the latter strategy achieves slower voltage recovery and regains slightly lower steady-state voltages. This example indicates that the SRLM has high potential in helping an off-the-shelf UVLS scheme select the most critical sites to achieve fast voltage recovery at the cost of curtailing loads as few as possible.

TABLE VI
VTSI PREDICTION PERFORMANCES IN RENEWABLE ENERGY CONTEXTS (WITH WIND POWER)

| Dataset | $MAE$ | $RMSE$ | $R_{el}/\%$ |
|---|---|---|---|
| Training set (9000) | 0.0126 | 0.0192 | 98.23 |
| Validation set (2000) | 0.0131 | 0.0195 | 98.00 |
| Testing set (2500) | 0.0133 | 0.0196 | 97.85 |

### *F. Applicability in Renewable Energy Contexts*

The test results reported in the above subsections involve SVS issues in the presence of conventional synchronous generators within the Nordic test system. In fact, with no limitation to the types of generators or loads, the proposed SRLM can also be applied to power systems with renewable power generation if the uncertainty and variability of renewable energy sources (RESs) are sufficiently considered during the process of offline VTSI dataset preparation. To demonstrate the applicability of the SRLM in renewable energy contexts, additional tests were carried out here by replacing some of the power plants in the Nordic test system with renewable wind farms. Specifically, the following generators in Fig. 5 were replaced by wind turbines: {g1, g2, g3, g5, g6, g7, g9, g10, g11, g12}. By doing so, RESs take up 35% of power generation in the system. Based on the test setup in existing studies on data-driven power system stability monitoring with RESs [48], the generation outputs of these wind turbines were set to randomly vary from 0 to their rated capacities, so as to account for the uncertainty and variability of RESs. Based on the case generation procedure described in Section IV-A, all the training, validation, and testing datasets for VTSI prediction were re-generated by incorporating these new settings related to RESs. Correspondingly, the SRLM was rebuilt, and its VTSI prediction performances on the newly generated datasets are presented in Table VI.

Clearly, the SRLM still achieves satisfactory performances in the context of high RES penetration. It maintains the VTSI estimation errors w.r.t. $MAE$ and $RMSE$ at relatively low levels of less than 0.015 and 0.02, respectively. With such a high precision, it manages to reliably identify more than 97.8% of the critical load buses for UVLS implementation. Hence, once the RES generation is considered in the offline procedure of learning data preparation, the proposed SRLM has high potential to adapt well to high renewables contexts. Nonetheless, how the penetration level of RESs in the system influences the performance of the proposed SRLM deserves more systematic investigations. Also, considering the significant volatility of RES generation in different months and seasons, advanced DL techniques like transfer learning may be introduced to further enhance the performance of the proposed SRLM.

## V. CONCLUSION

For the sake of adaptively selecting the most effective UVLS sites for power system online SVS control, this paper develops an intelligent SRLM-based scheme to predict short-term VTSI in a reliable and efficient way. By carefully incorporating the emerging GCN technique with the powerful LSTM algorithm, the SRLM is able to sufficiently learn the inherent structural and temporal features of a given power system from various SVS dynamics. Such a superior learning capability contributes to its high VTSI prediction precision during online application. Extensive numerical tests on the Nordic benchmark system demonstrate that the proposed scheme achieves much better VTSI prediction performances than conventional methods. Besides, it exhibits desirable robustness to practical imperfect conditions, e.g., unexpected structural errors/changes and PMU measurement noises. With the most critical sites for UVLS implementation reliably identified, it shows high potential in enhancing online SVS control. Note that the applicability of the SRLM is not limited to adaptive control site selection. In relevant future research, the SRLM can be extended to further optimize UVLS amounts for more systematic data-driven SVS control. Besides, how to work out more systematic solutions to address practical conditions like PMU data losses and high renewable energy penetration can be taken as future research directions.

## REFERENCES


[1] G. Bedi, G. K. Venayagamoorthy, R. Singh *et al.*, "Review of internet of things (iot) in electric power and energy systems," *IEEE Internet Things J.*, vol. 5, no. 2, pp. 847–870, 2018.

[2] L. Lin, X. Guan, Y. Peng *et al.*, "Deep reinforcement learning for economic dispatch of virtual power plant in internet of energy," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 6288–6301, 2020.

[3] C. Ren and Y. Xu, "A universal defense strategy for data-driven power system stability assessment models under adversarial examples," *IEEE Internet Things J.*, vol. 10, no. 9, pp. 7568–7576, 2023.

[4] T. Liu, Y. Song, L. Zhu, and D. J. Hill, "Stability and control of power grids," *Annu. Rev. Control Robot. Auton. Syst.*, vol. 5, pp. 689–716, 2022.

[5] C. Ren, Y. Xu, and R. Zhang, "An interpretable deep learning method for power system transient stability assessment via tree regularization," *IEEE Trans. Power Syst.*, vol. 37, no. 5, pp. 3359–3369, 2022.

[6] P. Kundur, J. Paserba, V. Ajjarapu *et al.*, "Definition and classification of power system stability," *IEEE Trans. Power Syst.*, vol. 19, no. 2, pp. 1387–1401, 2004.

[7] T. Van Cutsem and C. Vournas, *Voltage Stability of Electric Power Systems*. Norwell, MA, USA: Kluwer, 1998.

[8] M. Glavic, D. Novosel, E. Heredia *et al.*, "See it fast to keep calm: Real-time voltage control under stressed conditions," *IEEE Power and Energy Magazine*, vol. 10, no. 4, pp. 43–55, 2012.

[9] Q. Li, Y. Xu, and C. Ren, "A hierarchical data-driven method for event-based load shedding against fault-induced delayed voltage recovery in power systems," *IEEE Trans. Ind. Informat.*, vol. 17, no. 1, pp. 699–709, 2021.

[10] L. Zhu, W. Wen, Y. Qu, F. Shen, J. Li, Y. Song, and T. Liu, "Robust representation learning for power system short-term voltage stability assessment under diverse data loss conditions," *IEEE Trans. Neural. Netw. Learn. Syst.*, pp. 1–13, 2023, Early Access.

[11] 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, p. 118347, 2022.

[12] Q. Song *et al.*, "On credibility of adversarial examples against learning-based grid voltage stability assessment," *IEEE Trans. Dependable Secure Comput.*, pp. 1–14, 2022, Early Access.

[13] C. Ren, Y. Xu, J. Zhao, R. Zhang, and T. Wan, "A super-resolution perception-based incremental learning approach for power system voltage stability assessment with incomplete pmu measurements," *CSEE J. Power Energy Syst.*, vol. 8, no. 1, pp. 76–85, 2022.

[14] Y. Li, S. Zhang, Y. Li, J. Cao, and S. Jia, "Pmu measurements-based short-term voltage stability assessment of power systems via deep transfer learning," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–11, 2023.

[15] S. Cheng, Z. Yu, Y. Liu, and X. Zuo, "Power system transient stability assessment based on the multiple paralleled convolutional neural network and gated recurrent unit," *Protection and Control of Modern Power Systems*, vol. 7, no. 1, p. 39, 2023.

[16] L. Zhu, W. Wen, J. Li, and Y. Hu, "Integrated data-driven power system transient stability monitoring and enhancement," *IEEE Trans. Power Syst.*, pp. 1–13, 2023, Early Access.

[17] T. Van Cutsem and C. Vournas, "Emergency voltage stability controls: An overview," in *Proc. IEEE Power Engineering Society General Meeting*. IEEE, 2007, pp. 1–10.

[18] K. Mollah, M. Bahadornejad, N.-K. Nair, and G. Ancell, "Automatic under-voltage load shedding: A systematic review," in *2012 IEEE Power and Energy Society General Meeting*. IEEE, 2012, pp. 1–7.

[19] H. Bai and V. Ajjarapu, "A novel online load shedding strategy for mitigating fault-induced delayed voltage recovery," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 294–304, 2011.

[20] Y. Dong, X. Xie, K. Wang *et al.*, "An emergency-demand-response based under speed load shedding scheme to improve short-term voltage stability," *IEEE Trans. Power Syst.*, vol. 32, no. 5, pp. 3726–3735, 2017.

[21] H. Yang, N. Li, Z. Sun, D. Huang, D. Yang, G. Cai, C. Liu, T. Zhang, and W. Zhang, "Real-time adaptive uvls by optimized fuzzy controllers for short-term voltage stability control," *IEEE Trans. Power Syst.*, vol. 37, no. 2, pp. 1449–1460, 2022.

[22] L. Zhu, C. Lu, and Y. Luo, "Response-based adaptive under-voltage load shedding using oblique energy decision tree," in *IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia)*, 2019, pp. 3657–3662.

[23] J. Zhang, C. Lu, J. Si *et al.*, "Deep reinforcement leaming for short-term voltage control by dynamic load shedding in china southem power grid," in *Proc. Int. Joint Conf. Neural Networks*. IEEE, 2018, pp. 1–8.

[24] Q. Huang, R. Huang, W. Hao, J. Tan, R. Fan, and Z. Huang, "Adaptive power system emergency control using deep reinforcement learning," *IEEE Trans. Smart Grid*, vol. 11, no. 2, pp. 1171–1182, 2020.

[25] L. Zhu, D. J. Hill, and C. Lu, "Auto-starting semisupervised-learning-based identification of synchrophasor data anomalies," *IEEE Internet Things J.*, vol. 9, no. 15, pp. 13 651–13 663, 2022.

[26] L. Zhu and J. Lin, "Learning spatiotemporal correlations for missing noisy pmu data correction in smart grid," *IEEE Internet Things J.*, vol. 8, no. 9, pp. 7589–7599, 2021.

[27] Q. Li, Y. Xu, C. Ren, and R. Zhang, "A probabilistic data-driven method for response-based load shedding against fault-induced delayed voltage recovery in power systems," *IEEE Trans. Power Syst.*, vol. 38, no. 4, pp. 3491–3503, 2023.

[28] B. Sapkota *et al.*, "Dynamic var planning in a large power system using trajectory sensitivities," *IEEE Trans. Power Syst.*, vol. 25, no. 1, pp. 461–469, 2010.

[29] M. Paramasivam *et al.*, "Dynamic optimization based reactive power planning to mitigate slow voltage recovery and short term voltage instability," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 3865–3873, 2013.

[30] Y. Xu *et al.*, "Multi-objective dynamic var planning against short-term voltage instability using a decomposition-based evolutionary algorithm," *IEEE Trans. Power Syst.*, vol. 29, no. 6, pp. 2813–2822, 2014.

[31] I. A. Hiskens *et al.*, "Trajectory sensitivity analysis of hybrid systems," *IEEE Trans. Circuits Syst. I*, vol. 47, no. 2, pp. 204–220, 2000.

[32] M. Defferrard *et al.*, "Convolutional neural networks on graphs with fast localized spectral filtering," in *Advances in Neural Information Processing Systems*, 2016, pp. 3844–3852.

[33] T. N. Kipf *et al.*, "Semi-supervised classification with graph convolutional networks," in *Int. Conf. Learning Representations*, 2017, pp. 1–14.

[34] Y. Seo *et al.*, "Structured sequence modeling with graph convolutional recurrent networks," in *Int. Conf. Neural Information Processing*. Springer, 2018, pp. 362–373.

[35] Y. Song *et al.*, "On extension of effective resistance with application to graph laplacian definiteness and power network stability," *IEEE Trans. Circuits Syst. I*, vol. 66, no. 11, pp. 4415–4428, 2019.

[36] J. Yu, D. J. Hill, V. O. Li, and Y. Hou, "Synchrophasor recovery and prediction: A graph-based deep learning approach," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 7348–7359, 2019.

[37] B. Chen, Q. Wu, M. Li, and K. Xiahou, "Detection of false data injection attacks on power systems using graph edge-conditioned convolutional networks," *Protection and Control of Modern Power Systems*, vol. 8, no. 1, pp. 1–12, 2023.

[38] Y. Luo, C. Lu, L. Zhu, and J. Song, "Graph convolutional network-based interpretable machine learning scheme in smart grids," *IEEE Trans. Autom. Sci. Eng.*, vol. 20, no. 1, pp. 47–58, 2023.

[39] W. Xia, D. He, and L. Yu, "Locational detection of false data injection attacks in smart grids: A graph convolutional attention network approach," *IEEE Internet Things J.*, 2023, Early Access.

[40] J. Huang, L. Guan, Y. Su, H. Yao, M. Guo, and Z. Zhong, "System-scale-free transient contingency screening scheme based on steady-state information: A pooling-ensemble multi-graph learning approach," *IEEE Trans. Power Syst.*, vol. 37, no. 1, pp. 294–305, 2022.

[41] D. I. Shuman *et al.*, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 83–98, 2013.

[42] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT press Cambridge, 2016.

[43] T. Van Cutsem and L. Papangelis, "Description, modeling and simulation results of a test system for voltage stability analysis," Tech. Rep., 2013.

[44] T. Van Cutsem *et al.*, "Test systems for voltage stability studies: Ieee task force on test systems for voltage stability analysis and security assessment," *IEEE Trans. Power Syst.*, vol. 35, no. 5, pp. 4078–4087, 2020.

[45] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-c. Woo, "Convolutional lstm network: A machine learning approach for precipitation nowcasting," in *Advances in Neural Information Processing Systems*, 2015, pp. 1–9.

[46] L. Zhu, C. Lu, and D. J. Hill, "Hierarchical deep learning machine for power system online transient stability prediction," *IEEE Trans. Power Syst.*, vol. 35, no. 3, pp. 2399–2411, 2020s.

[47] *IEEE Standard for Synchrophasor Measurements for Power Systems*, IEEE Std. C37.118.1-2011, 2011.

[48] Y. 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 Trans. Smart Grid*, vol. 10, no. 5, pp. 5663–5674, 2019.