System Strength Assessment Based on Multi-task Learning

Baoluo Li, Shiyun Xu, *Member*, *CSEE*, Huadong Sun, *Fellow*, *CSEE*, Zonghan Li, and Lin Yu

***Abstract*—Increase in permeability of renewable energy sources (RESs) leads to the prominent problem of voltage stability in power system, so it is urgent to have a system strength evaluation method with both accuracy and practicability to control its access scale within a reasonable range. Therefore, a hybrid intelligence enhancement method is proposed by combining the advantages of mechanism method and data driven method. First, calculation of critical short circuit ratio (CSCR) is set as the direction of intelligent enhancement by taking the multiple renewable energy station short circuit ratio as the quantitative indicator. Then, the construction process of CSCR dataset is proposed, and a batch simulation program of samples is developed accordingly, which provides a data basis for subsequent research. Finally, a multi-task learning model based on progressive layered extraction is used to simultaneously predict CSCR of each RESs connection point, which significantly reduces evaluation error caused by weak links. Predictive performance and anti-noise performance of the proposed method are verified on the CEPRI-FS-102 bus system, which provides strong technical support for real-time monitoring of system strength.**

***Index Terms*—Critical short circuit ratio, hybrid intelligence enhancement, multi-task learning, system strength.**

I. Introduction

WITH rapid development of renewable energy sources (RESs), mainly wind power and photovoltaic, stable operation and planning of power system face new challenges [1], [2]. On one hand, reduction in the proportion of conventional generation leads to a weakening of power system strength, which in turn limits the RESs integration scale [3]. On the other hand, as the proportion of RESs in system increases, the scenario of RESs integration into a weak AC system frequently occurs, leading to wide-frequency oscillation, transient overvoltage and other problems [4], [5]. Therefore, a system strength evaluation (SSE) method, which does not rely on off-line simulation and only uses response information, is urgently needed to ensure stable operation of the power system and solve limitation of RESs integration scale [6]–[8].

In order to effectively quantify impact of RESs integration scale on system strength, in [9], a multiple renewable energy station short circuit ratio (MRSCR) is constructed based on the physical nature of short circuit ratio (SCR). MRSCR fully considers interaction between the RESs plants and reactive power influence of the RESs power generation equipment, so it has good characterization ability. In the setting of stability criterion, the critical value of MRSCR, namely critical short circuit ratio (CSCR), is obtained by calculating typical grid parameters. SSE is accomplished by comparing MRSCR and CSCR. However, CSCR changes dynamically under different scenarios, and the CSCR set by combining engineering experience with simulation, can only roughly characterize the critical state of the system. Therefore, accuracy of SSE method based on MRSCR is at a low level, which cannot meet current power grid assessment needs. In fact, most SCR indicators suitable for RESs use the above method to set the corresponding CSCR, such as weighted short circuit ratio (WSCR) [10], composite short circuit ratio (C-SCR) [11], equivalent circuit-based short circuit ratio (ECSCR) [12].

In addition to the above method, some studies construct SCR with a clear threshold according to voltage stability conditions, to provide a basis for calculation of CSCR. Reference [13] constructed SCR-S based on ratio of system short-circuit capacity and RESs grid-connected capacity, and proposed calculation expression of CSCR based on power transmission limit. In [14], site-dependent short circuit ratio (SDSCR) was defined, and the characteristic value of static voltage critical stability was derived as CSCR. In [15], a generalized short circuit ratio (GSCR) is defined from a small disturbance stability Angle, and it is proposed when Hopf bifurcation occurs in the system, GSCR is CSCR. The above calculation methods have strict mechanism support, but there are many assumptions in the derivation process. Once all assumptions are unshackled, its usefulness may decline. In addition, calculation parameters required by some methods are difficult to be measured in real time. CSCR of the above SCR indices are listed in Table I, where S_{ac} and $S_{\max}$ are short circuit capacity and maximum transmission capacity, respectively. Calculation of CSCR in Table I relies on manual experience or assumptions, is not a response-based method, and assessment accuracy cannot meet requirements.

In recent years, due to the breakthrough development of deep learning (DL), application of DL to analyze power system stability has once again become a research hotspot [16]–[20]. This research directly uses the DL model to identify

Manuscript received January 27, 2023; revised March 27, 2023; accepted April 28, 2023. Date of online publication December 28, 2023; date of current version December 31, 2023.

B. L. Li is with the Key Laboratory of Power System Intelligent Dispatch and Control of Ministry of Education, Shandong University, Jinan 250014, China.

S. Y. Xu, H. D. Sun (corresponding author, email: sunhd@epri.sgcc.com.cn), Z. H. Li and L. Yu are with China Electric Power Research Institute, Beijing 100192, China.

DOI: 10.17775/CSEEJPES.2023.00440

TABLE I
HYPER-PARAMETER SETTING OF MTL MODEL

Indicator name	Numerical setting	System strength	Attribute
MRSCR [9]	CSCR < 2	weak system	empirical value
WSCR [10]	CSCR < 1.5	weak system	empirical value
C-SCR [11]	CSCR < 1.5	weak system	empirical value
ECSCR [12]	CSCR < 3	weak system	empirical value
SCR-S [13]	CSCR $= S_{\mathrm{ac}}/S_{\max}$	critical stability	theoretical value
SDSCR [14]	CSCR $= 1$	critical stability	theoretical value
GSCR [15]	eigenvalue $= 0$	critical stability	theoretical value

the stable form of the system, which has advantages of fast identification speed and flexible modeling methods. Obviously, introducing this research paradigm into the field of SSE can not only simplify the evaluation process, but also avoid defects of the CSCR calculation method. However, limited by shortcomings of DL in terms of generalization performance and interpretability, relying solely on DL models to solve stability problems will face security risks. Therefore, this paper focuses on the mechanism method, and regards DL as an auxiliary decision-making tool, thus forming a hybrid intelligence enhancement method. First, considering MRSCR has strong characterization ability, this indicator is selected as the research object of intelligence enhancement. By analyzing the evaluation process of MRSCR, it is determined calculation of CSCR is the link to be enhanced. Then, according to the physical meaning of CSCR, the CSCR dataset corresponding to MRSCR is constructed. Finally, considering the large number of RESs connection points, a multi-task learning (MTL) model based on progressive layered extraction (PLE) is used to mine the mapping relationship between power flow and CSCR of each RESs connection point. Effectiveness of the proposed method are demonstrated on the CEPRI-FS-102 bus system, and compared with various DL models and empirical values.

The rest of this paper is organized as follows. In Section II, the principle of MRSCR and construction process of CSCR dataset are introduced. In Section III, necessity of multi-task learning is analyzed and principle of PLE are introduced. In Section IV, effectiveness of the proposed method is demonstrated. In Section V, conclusions are given.

II. SCR SELECTION AND CSCR DATASET CONSTRUCTION

A. Definition and Expression of MRSCR

Assuming current injected into the AC system by the bus of each RESs connection point is $\dot{I}_1, \dot{I}_2, \cdots, \dot{I}_n$, then bus voltage $\dot{U}_{\mathrm{RE}1}, \dot{U}_{\mathrm{RE}2}, \cdots, \dot{U}_{\mathrm{RE}n}$ of each connection point can be expressed as:

$$\begin{bmatrix}\dot{U}_{\mathrm{RE}1}\\ \dot{U}_{\mathrm{RE}2}\\ \vdots \\ \dot{U}_{\mathrm{RE}n}\end{bmatrix} = \begin{bmatrix}\dot{Z}_{\mathrm{eq}11} & \dot{Z}_{\mathrm{eq}12} & \cdots & \dot{Z}_{\mathrm{eq}1n}\\ \dot{Z}_{\mathrm{eq}21} & \dot{Z}_{\mathrm{eq}22} & \cdots & \dot{Z}_{\mathrm{eq}2n}\\ \vdots & \vdots & \ddots & \vdots \\ \dot{Z}_{\mathrm{eq}n1} & \dot{Z}_{\mathrm{eq}n1} & \cdots & \dot{Z}_{\mathrm{eq}nn}\end{bmatrix}\begin{bmatrix}\dot{I}_1\\ \dot{I}_2\\ \vdots \\ \dot{I}_n\end{bmatrix} \tag{1}$$

where $\dot{Z}_{\mathrm{eq}ij}$ is element in row i and column j of equivalent impedance matrix of the AC network at RESs connection point. MRSCR is defined as relative magnitude between system nominal voltage and RESs generated voltage, as follows:

$$\begin{aligned}\mathrm{MRSCR}_i &= \frac{|\dot{U}_{\mathrm{N}i}|}{|\dot{U}_{\mathrm{RE}i}|} = \frac{|\dot{U}_{\mathrm{N}i}|}{\left|\dot{Z}_{\mathrm{eq}ii}\dot{I}_i + \sum_{j=1,j\neq i}^{n}\dot{Z}_{\mathrm{eq}ij}\dot{I}_j\right|}\\ &= \left|\frac{\dot{U}_i^*\dot{U}_{\mathrm{N}i}}{\dot{Z}_{\mathrm{eq}ii}}\right| \Big/ \left|\dot{S}_{\mathrm{RE}i} + \sum_{j=1,j\neq i}^{n}\frac{\dot{Z}_{\mathrm{eq}ij}\dot{U}_i^*}{\dot{Z}_{\mathrm{eq}ii}\dot{U}_j^*}\dot{S}_{\mathrm{RE}j}\right|\end{aligned} \tag{2}$$

where $\dot{U}_{\mathrm{N}i}$ is bus nominal voltage of i-th RESs connection point; $\dot{U}_{\mathrm{RE}i}$ is voltage generated by RESs at i-th bus; $\dot{S}_{\mathrm{RE}i}$ is apparent power injected by bus of i-th RESs connection point.

In order to consider engineering application value, index selects CSCR of connection point of RESs power generation equipment as 1.5 according to calculation results of typical power grid parameters. However, the empirical value obtained by qualitative analysis is difficult to accurately represent critical state of the system, so it is necessary to realize the quantitative analysis of CSCR by intelligent enhancement, so as to improve adaptability of this indicator.

B. Construction Process of CSCR Dataset

System strength requires explicit stability criteria to be assessed. In analysis of static voltage stability, it is considered the power system reaches transmission power limit as the critical state of the system strength, that is, stability criterion [21]. In this state, the MRSCR calculated by equation (2) is CSCR. CSCR is the dividing line between whether a power flow has a solution, and it can establish the relationship between MRSCR and static voltage stability.

According to the above physical significance, construction steps of CSCR dataset is as follows:

1) Total power of RESs is randomly divided into active power of each station, and Initial power flow state (IPFS) of system is obtained.

2) In order to obtain maximum available power (MAP) of each RESs connection point, continuation power flow (CPF) method is repeatedly used to trace power-voltage curve. First, select an RESs power plant and make its active power increase according to constant power factor, in which surplus power is absorbed by the generator of the slack bus. Then, steady-state behavior of the system under the power variation is tracked until MAP of the corresponding connection points is obtained.

3) CSCR of each connection point is calculated according to the MAP as the prediction label, and electrical features under current IPFS are extracted as input features.

4) According to the required number of samples, repeat steps 1–3 to obtain a large number of point sets of static voltage stability critical point. Each point set constitutes the boundary of static voltage stability domain corresponding to the connection point. To facilitate storage of the dataset, input features and prediction labels of all samples are combined in a two-dimensional array.

Considering the ways in which active power of multiple RESs power plants increase simultaneously are complex and varied, step 2 obtains a relatively conservative CSCR by simulating active power increase of a single plant. To facilitate understanding of the above process, it is drawn in Fig. 1, where M is the number of features, N is the number of samples,

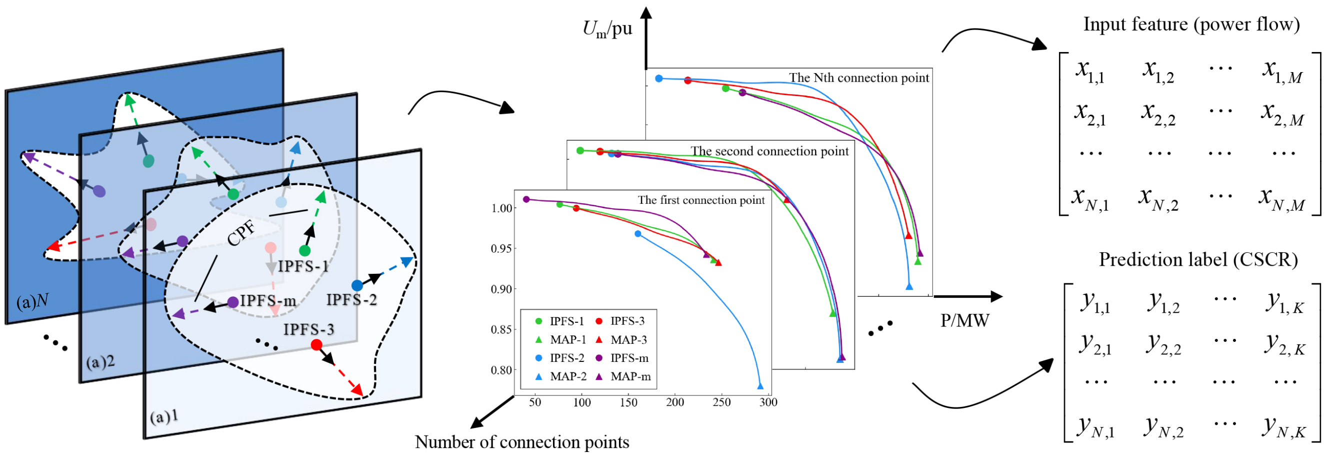

Fig. 1. Construction process and data structure of CSCR dataset.

and K is the number of connection points. (a)1–(a)N, each plane represents the stable domain boundary corresponding to different RESs connection points.

III. Intelligent Enhancement Method Based on MTL

A. Principle of MTL

Typical learning mode of DL is single-task learning (STL), which uses a specific model to solve a single task. For complex multi-tasks, it can also be decomposed into simple and independent single tasks to solve. When STL is used for CSCR prediction tasks, the training process is shown in Fig. 2(a).

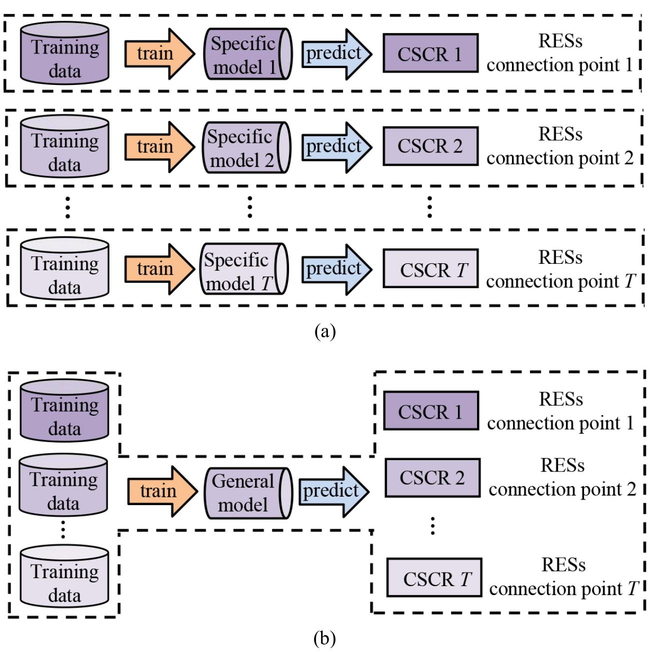

Fig. 2. Comparison between single-task learning and multi-task learning. (a) Single-task learning. (b) Multi-task learning.

Figure 2(a) shows STL needs to train T specific models for T RESs connection points. However, a power system with high penetration of RESs has a large number of RESs connection points, which will lead to increased training complexity and maintenance cost of the model. In addition, according to electrical distance, there is a strong coupling relationship between maps of different union points, while STL ignores correlation information between each union point. Therefore, STL is difficult to apply to CSCR prediction task, and MTL needs to be used to solve the above problems.

MTL is a learning mode proposed for multi-task scenarios. Its essence is to compute multiple tasks in parallel by a general model with a sharing mechanism. According to feature sharing mode, MTL is divided into hard parameter sharing and soft parameter sharing. The former is multiple tasks use the same feature sharing layer, while the latter is each task has dedicated feature parameters, which need to be regularized to achieve the effect of sharing information. Due to the large number of RESs connection points, hard parameter sharing is simpler in terms of model structure and feature parameters. Therefore, hard parameter sharing is selected for CSCR prediction task, and the training process is shown in Fig. 2(b). The general model in Fig. 2(b) can obtain prediction results of all connection points at one time, and its loss function is as follows:

$$\min \left[\sum_{k=1}^{K} \hat{E}_k(w_s, w_k) \right] \tag{3}$$

where $\hat{E}_k(w_s, w_k) = \frac{1}{N}\sum_{i=1}^{N} E(f_t(x_i, w_s, w_k), y_i^k)$ is the prediction error of k-th connection point; K is the number of RESs connection points; x_i is the i-th sample; y_i^k is the true value of k-th connection point in i-th sample; w_s and w_k are shared parameters and task parameters among different tasks, respectively. f is the prediction function. Equation (3) shows MTL helps the model to extract better abstract features by mining shared information, thus improving prediction accuracy of each task.

B. Principle of PLE

PLE is a MTL model based on hard parameter sharing mechanism, consisting of shared network, expert network, gating network, and tower network [22]. Taking two connection points as an example, the model structure of PLE is shown in Fig. 3, where the lower corner mark of the sub-network

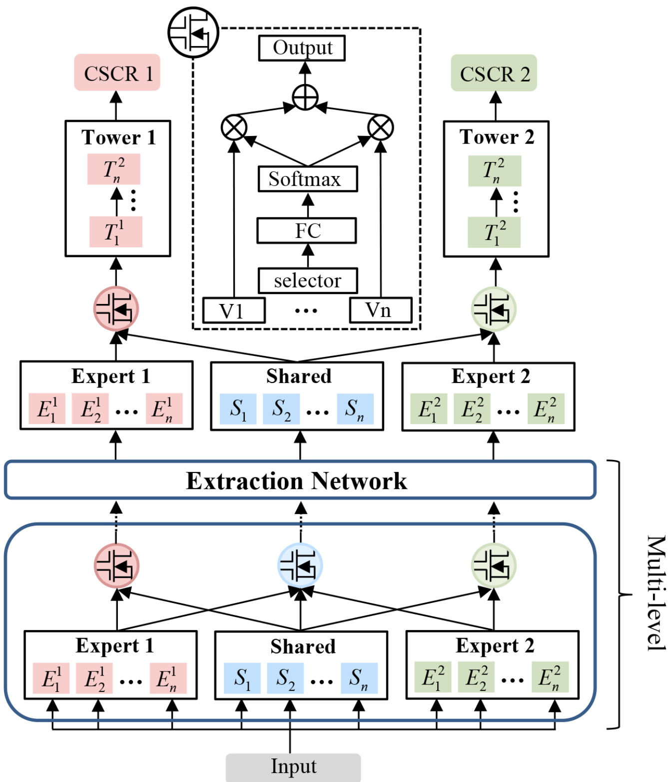

Fig. 3. The model structure of PLE.

represents the number of tasks, and the upper corner mark represents the number of networks or layers.

Workflow of the PLE in Fig. 3 is as follows:

1) According to number of RESs connection points, the extraction network layer composed of shared network and different expert networks is constructed. Both sharing network and expert network consist of multiple sub-networks. Deeper feature extraction can be realized by setting up multi-layer extraction networks.

2) Input features are extracted by sharing network and expert network first. Then, the gating network selectively fuses the extracted features. The gating network is a single-layer feedforward network, and its structure is shown in the middle part of Fig. 3. V in Fig. 3 represents input vector, and gating network output formula of task k is:

$$g^k(x) = w^k(x)S^k(x) \tag{4}$$

where x is input, $w^k(x)$ is weighting function of k-th task:

$$w^k(x) = \text{Softmax}(w_g^k x) \tag{5}$$

where w_g^k is parameter matrix. $S^k(x)$ is a selected matrix consisting of output of the shared network and expert network for task k:

$$S^k(x) = \left[(E_1^k)^{\mathrm{T}}, \cdots, (E_n^k)^{\mathrm{T}}, (S_1^k)^{\mathrm{T}}, \cdots, (S_n^k)^{\mathrm{T}}\right]^{\mathrm{T}} \tag{6}$$

3) Output of gating network enters each tower network to obtain the predicted output of each connection point. Calculation method of the tower network is:

$$y^k(x) = t^k(g^k(x)) \tag{7}$$

where t^k is tower network of task k.

Compared with traditional MTL model, PLE explicitly separates the sharing network from the expert network, so the expert network of different tasks can concentrate on learning different knowledge. Combined with dynamic fusion of gated networks, the tower network effectively handles balance between dependencies of different connection points.

C. Process of Intelligent Enhancement Method

According to the timing relationship, the process of the

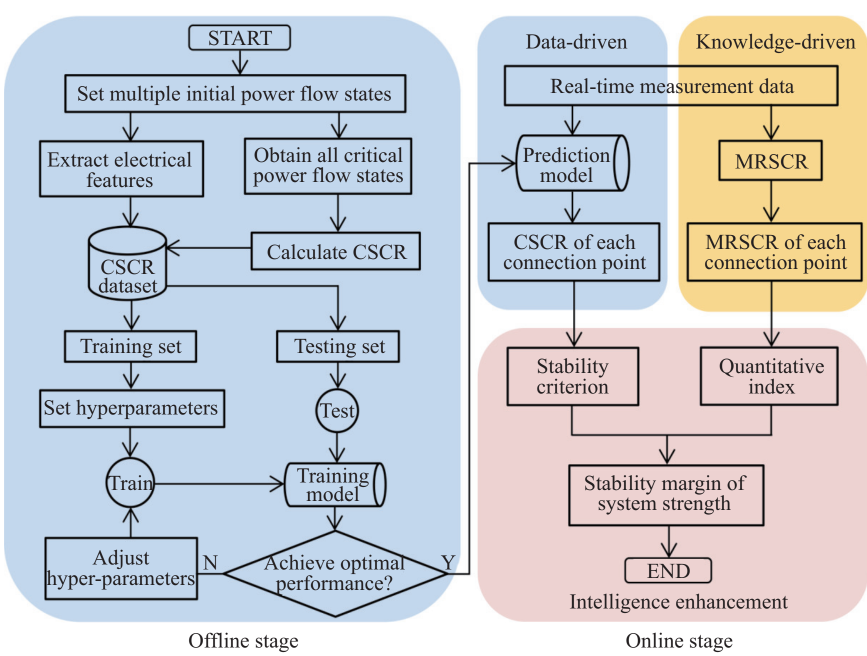

Fig. 4. The evaluation process of the proposed method.

proposed method is divided into offline stage and online stage, as shown in Fig. 4.

Offline stage: First, the same number of IPFS are randomly generated according to required number of samples, and corresponding critical power flow state is obtained through CPF. Electrical features of IPFS and the CSCR of each connection point are used to construct the dataset. Then, the dataset is divided into a training set and a testing set, and hyper-parameters such as number of tasks, number of sub-networks and the network structure are set. Finally, the mapping relationship between electrical feature and CSCR is constructed using PLE on the training set. According to test results of the testing set, hyper-parameters are adjusted repeatedly until evaluation performance is optimal.

Online stage: First, real-time measurement data is input into the prediction model, and the model quickly predicts the CSCR of each connection point under the current state according to learned mapping relationship. Meanwhile, MRSCR of each connection point is calculated according to (2). Then, margin analysis of system strength can be completed by calculating the difference between MRSCR and CSCR.

Figure 4 shows the proposed method completes SSE by calculating quantitative indicator and prediction stability criterion, which has the dual driving characteristics of knowledge and data. The method uses the complementary relationship between mechanism method and data-driven method to achieve intelligent enhancement, to improve practicability and accuracy of SSE method.

IV. Case Study

Power system simulation software is PSD-BPA and deep learning framework is Pytorch. Computer is configured with an AMD Ryzen 7 5800 8-core 3.40 GHz CPU, 16 GB RAM and NVIDIA GeForce RTX3060.

A. CSCR Dataset

In this experiment, CEPRI-FS-102 bus system is used as test system, and its topology is shown in Fig. 5. The system has 102 buses in total, of which active power of RESs is 900 MW (all located at sending end), and active power of conventional power generation is 3700 MW. The system contains 12 RESs power plants, including 6 wind power plants and 6 photovoltaic power plants.

In order to realize the construction process of CSCR dataset, a batch simulation program based on Python and PSD-BPA is developed. The program has the following functions: 1) Randomly generate a specified number of IPFSs, eliminate the flow does not converge; 2) Under each IPFS, CPF is carried out for each RESs connection point successively, and MAP is automatically located; 3) Extract required electrical features and calculate prediction label.

For test system, 10,000 IPFS are randomly generated by the program, and voltage amplitudes and phase angles of all buses, active power and reactive power of lines are extracted as input features. Meanwhile, to improve difficulty of prediction tasks, the CSCR of the RESs device connection point is used as the prediction label to increase number of tasks. In addition, fluctuation range of active power of W_2 and other RESs plants is set as 0–200 MW and 10–200 MW, respectively, to increase difference of label distribution. After dataset is generated according to the above settings, distribution of labels is counted, as shown in Fig. 6. CSCR of each RESs connection point in Fig. 6 changes dynamically, and fluctuation range is between 1.2 and 2.3. Obviously, severe fluctuation makes it difficult for experience value 1.5 to adapt to various scenarios, which leads to large errors in evaluation results. In addition, from distribution of CSCR, it can be seen there are both similarities and differences among all connection points. In particular, as shown in Fig. 6(b), distribution of peaks occurs between 1.2 and 1.3. Therefore, this dataset is suitable for verifying feasibility of intelligence enhancement and rationality of applying MTL.

B. Performance Comparison of Different Models

In order to verify effectiveness of PLE, multiple models are selected for comparison in this experiment. In terms of STL, deep neural network (DNN), decision tree (DT), gradient

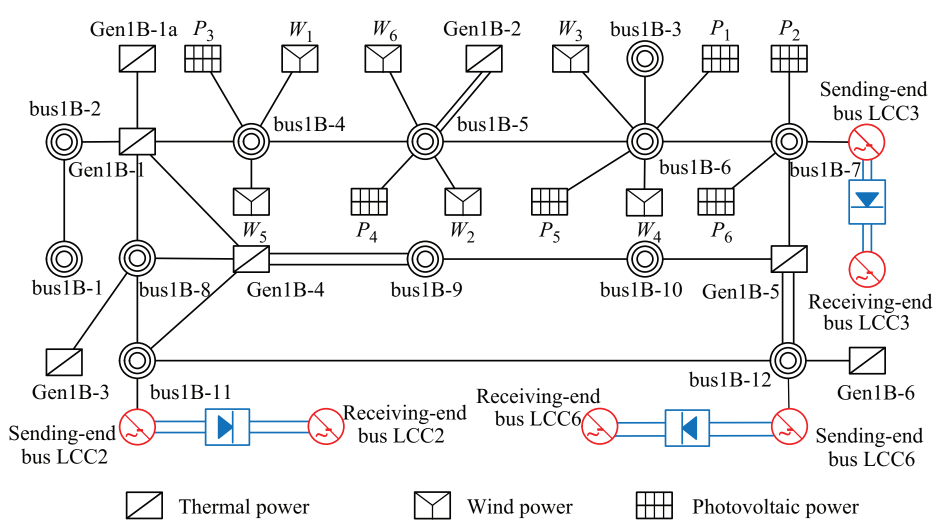

Fig. 5. CEPRI-FS-102 bus system topology.

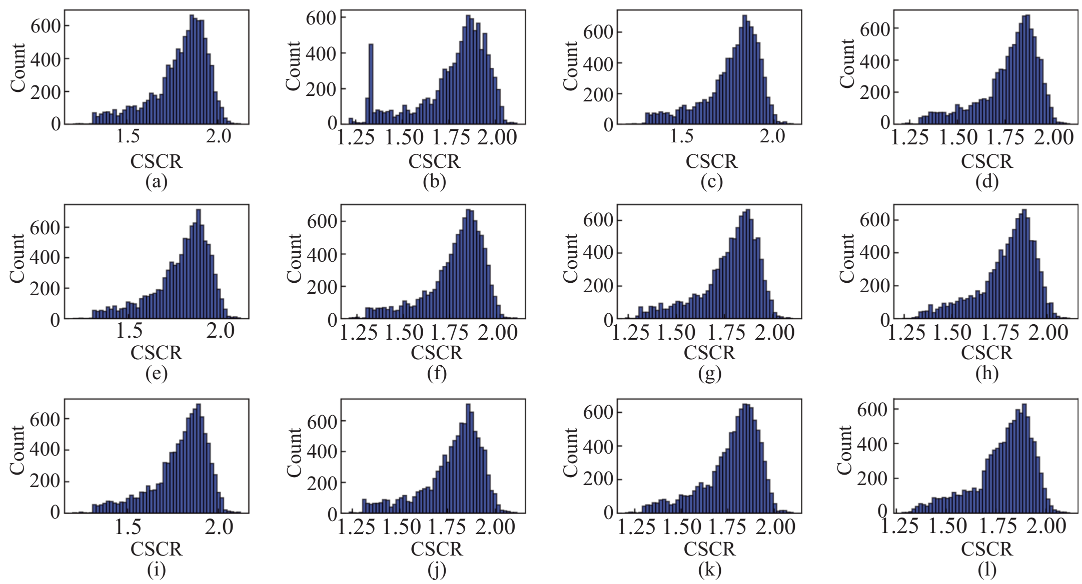

Fig. 6. The evaluation process of the proposed method. (a) Connection point of W_1. (b) Connection point of W_2. (c) Connection point of W_3. (d) Connection point of W_4. (e) Connection point of W_5. (f) Connection point of W_6. (g) Connection point of W_1. (h) Connection point of W_2. (i) Connection point of W_3. (j) Connection point of W_4. (k) Connection point of W_5. (l) Connection point of W_6.

boosting decision tree (GBDT) and support vector machine (SVM) as comparison model. Parameters of the above model are set as follows: he network structure of DNN is 512-256-128-64-1; maximum depth of DT and GBDT is 10, and number of GBDT base models is 50. Penalty factor C and nuclear parameter γ of SVM are 200 and 1. In terms of MTL, Bottom-Shared, One-gate Mixture-of-Experts (OMOE) and Multi-gate Mixture-of-Experts (MMOE) are selected as comparison models. Bottom-Shared consists only of shared networks and tower networks. On the basis of Bottom-Shared, OMOE and MMOE transform shared network into multiple expert networks and introduce gated units to capture task differences [21]. Hyper-parameters of the MTL model are shown in Table II.

TABLE II
HYPER-PARAMETER SETTING OF MTL MODEL

Hyper-parameter	Hyper-parameter value
Number of tasks	12
Optimizer	Adam
Loss function	MSE
Structure of expert network	512-256
Structure of shared network	512-256
Structure of tower network	128-64-1
Dropout	0.3
Number of sub-networks	4
Initial learning rate	0.001
Training epoch	50
Activation function	RELU

In order to comprehensively evaluate prediction performance of each model, mean absolute error (MAE), root mean Square Error (RMSE), mean absolute percentage error (MAPE) and R-square (R2) were selected as error indicators. The calculation equation of the above indicators is as follows:

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|(y_i - \hat{y}_i)| \tag{8}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{9}$$

$$\mathrm{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \tag{10}$$

$$\mathrm{R}^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y}_i)^2} \tag{11}$$

where y_i is true value, $\hat{y}_i$ is predicted value, and $\bar{y}_i$ is average value of y_i.

In addition to the above Settings, 80% of the CSCR dataset is divided into training set and the remaining 20% into testing set. Average of 20 experimental results is taken as final result, as shown in Fig. 7.

It can be seen from Fig. 7, SVM and DT have simple structures and relatively low prediction accuracy. GBDT and DNN have achieved better prediction results by virtue of ensemble learning and strong feature extraction capability, respectively. In terms of MTL, Bottom-shared, OMOE and MMOE fall into a seesaw phenomenon due to the complex correlation among 12 connection points, that is, prediction accuracy of some junction points is improved at the expense of prediction accuracy of other connection points. This phenomenon leads to a large difference in prediction accuracy among connection points. Compared with other MTL models, PLE has a higher balance among all connection points, and its error indicators are in a good state. This is because PLE explicitly separates shared experts from task-specific expert networks to avoid interference with invalid parameters. Second, PLE introduces multi-level expert and gated network to extract more abstract features and realize more efficient knowledge transfer between complex associative tasks. In addition, to further compare applicability of MTL, time cost of each method is calculated, as shown in Table III.

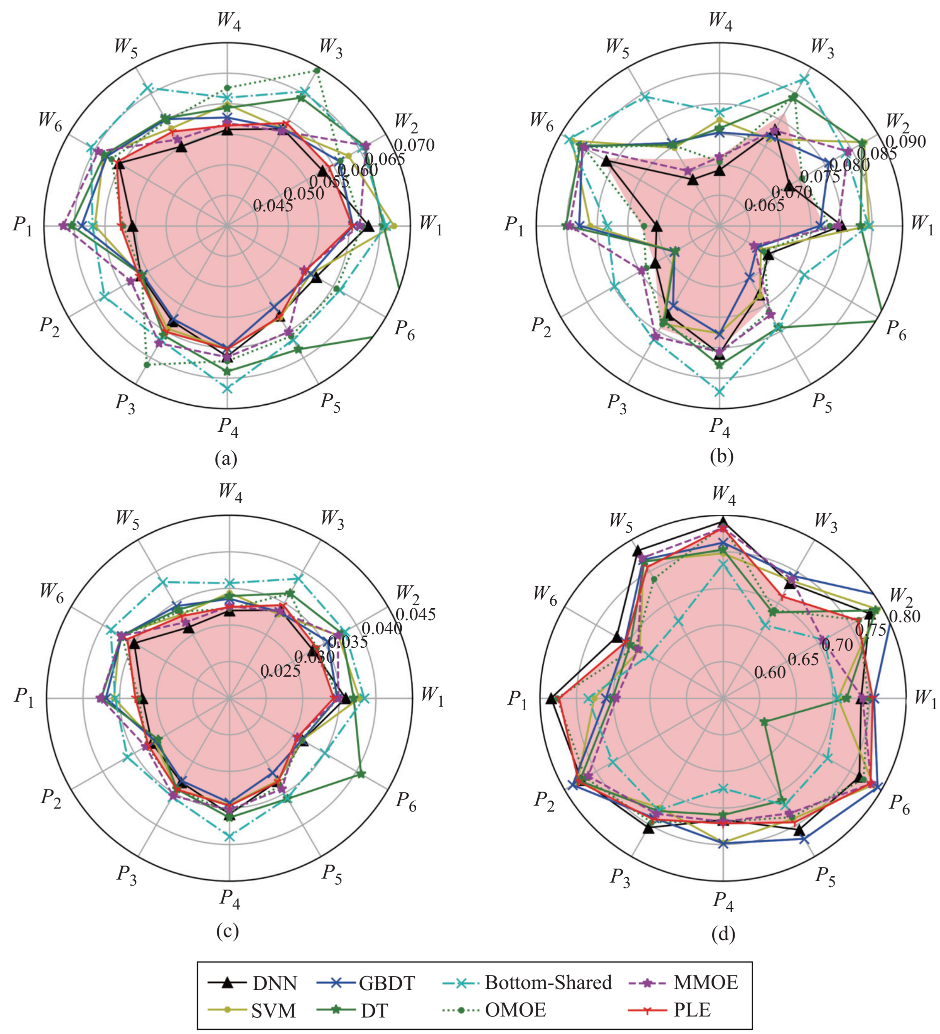

Fig. 7. Comparison of evaluation performance of each model. (a) MAE. (b) RMSE. (c) MAPE. (d) R2.

TABLE III
TIME COST OF EACH MODEL

Mode	Model	Training time (s)	Prediction time (s)
Offline simulation	CPF	—	7.46
STL	DT	10.63	0.00000437
	SVM	41.43	0.000312
	GBDT	246.79	0.0000324
	DNN	263.84	0.000319
MTL	Bottom-shared	107.98	0.0000929
	OMOE	132.36	0.000119
	MMOE	149.55	0.000123
	PLE	188.29	0.000156

As shown in Table III, offline simulation requires repeated use of CPF, so its prediction time cannot meet the requirements of real-time evaluation. For STL, DT and SVM as individual learners have a good time advantage. Training time of GBDT and DNN is relatively high due to accumulation mode and complex structure of the models. Once scale of RESs continues to expand, training time of both will continue to increase. It is obvious that MTL model is superior to STL model and offline simulation in terms of prediction accuracy and time cost. In particular, PLE has the advantages of prediction accuracy and time, so it has stronger engineering application value.

C. Effect Analysis of Intelligent Enhancement

In order to show the effect of intelligence enhancement, this experiment analyzes errors before and after enhancement and adverse consequences caused by errors. First, taking connection point W_1 power plant as an example, the training set and testing set are arranged in ascending order according to real value of the sample, and corresponding predicted value and empirical value are drawn in Fig. 8.

Figure 8 shows the predicted value of PLE is close to real value, while empirical value is far from real value of most samples. Obviously, empirical value is the main factor leading to evaluation error, so it is reasonable to apply DL technology to enhance calculation link of CSCR. In order to further quantify the promotion degree of predicted value to SSE, the following indicators are defined:

$$\begin{cases} E_c = \frac{1}{N_c} \sum_{i=1}^{N_c} \hat{y}_i - y_i, & \hat{y}_i > y_i \\ E_r = \frac{1}{N_r} \sum_{i=1}^{N_r} y_i - \hat{y}_i, & \hat{y}_i < y_i \end{cases} \tag{12}$$

where N_c is number of samples with predicted or empirical value greater than true value in the dataset, and N_r is number of samples with predicted or empirical value less than true value in the dataset. If predicted or empirical value is greater

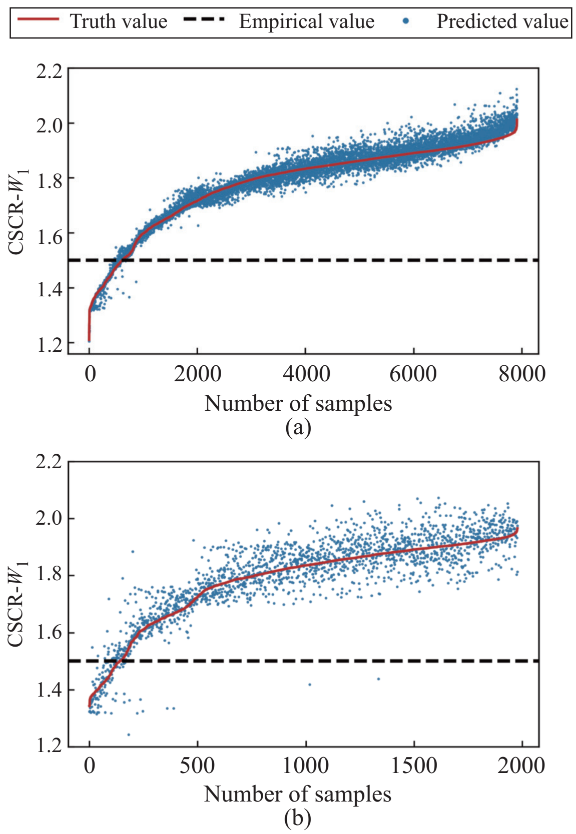

Fig. 8. Visual comparison of predicted and empirical value. (a) Training set. (b) Testing set.

than true value, it will be difficult for the RESs connection point to reach maximum transmission capacity, so E_c is a conservative error. On the contrary, the RESs access scale may exceed the MAP of the connection point, which leads to difficulty of maintaining the system voltage within normal range and the phenomenon of static voltage instability, so E_r is radical error. Four scenarios are defined based on E_c and E_r, as shown in Table IV.

TABLE IV
ERROR INDICATORS OF DIFFERENT SCENARIOS

Scenario	Dataset	Error indicator
C_1	Testing set	E_c
C_2	Training set	E_c
C_3	Testing set	E_r
C_4	Training set	E_r

According to Table IV, errors of predicted and empirical values in different scenarios are quantified, as shown in Fig. 9.

Both E_c and Er of empirical values in Fig. 9 are at a high level, which indicates the SSE method based on MRSCR cannot accurately evaluate system strength. More importantly, E_r is much larger than E_c, and empirical value is not conducive to stable operation of the system. In order to comprehensively quantify the enhancement effect, average error of 12 connection points in testing set is calculated. Compared with empirical value, mean values of E_r and E_c after enhancement decreased by 85.41% and 56.77%, respectively. This proves accuracy of SSE method can be effectively improved through

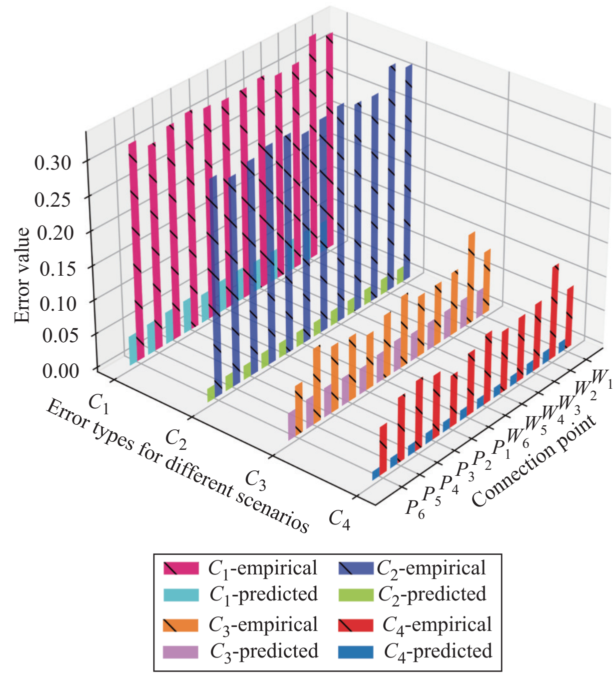

Fig. 9. Error comparison between predicted value and empirical value.

intelligent enhancement, especially significant reduction of E_r is helpful to ensure stable operation of power system.

D. Validity Analysis of PLE

In order to reveal the working principle of PLE, this experiment analyzes utilization of expert network, that is, weight assigned by gated network to each expert sub-network. The difference of this weight represents the freedom of choosing shared feature in tower network. MMOE with 4 expert sub-networks is compared with PLE with 2 expert sub-networks and 2 shared sub-networks. Output vector of the gated network was extracted and weight distribution is obtained using the Softmax function as shown in Fig. 10, where the sum of the weights of sub-networks is 1, and column height and error bar represent mean and standard deviation of the weights, respectively.

Results in Fig. 10 show PLE groups expert sub-networks with significantly different weights, while MMOE groups expert sub-networks with similar weights. This shows that PLE can provide strongly correlated shared features for tower networks by showing that separating shared and expert networks. Structure of PLE implements a more effective adaptive weighting method, which is helpful for mining the coupling relationship between union points, and is more suitable for synchronization prediction task of CSCR.

E. Anti-noise Performance Test of PLE

In order to verify anti-noise performance of PLE, this experiment sets up two kinds of noise for testing. First is to add 3% to 5% of measured value to test set as standard deviation of uniform noise, which is called type I noise. Second type of noise is burst noise, called type I noise. At any given sampling instant, there is a 20% probability the measured value will be noisy. This noise reflects random mutations in measurements

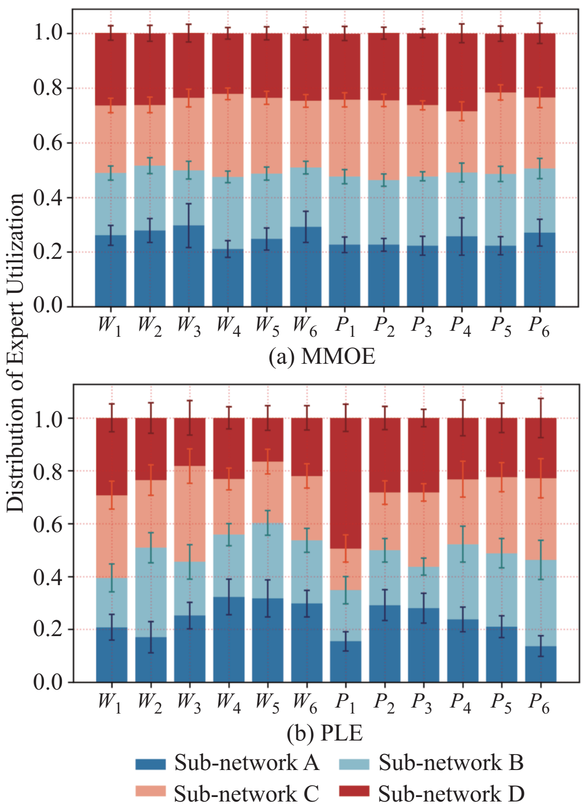

Fig. 10. Error comparison between predicted value and empirical value.

and is more realistic than uniform noise [24]. Type II noise is to add 3% to 5% of measured value as standard deviation of uniform noise to test set with a certain probability. 2 types of noise are added as follows:

$$X_{\text{noise}} = \begin{cases} X_{\text{ori}} \times (1 + K \times \theta), & \text{Type I} \\ X_{\text{ori}} \times (1 + K \times P \times \theta), & \text{Type II} \end{cases} \tag{13}$$

where X_{ori} is the original dataset, X_{noise} is the dataset containing noise, K is noise level, θ is Gaussian white noise whose mean value is 1 and variance is 0. A random number where P is 0 or 1, there is an 80% chance P is 0 and a 20% chance P is 1.

Considering randomness of the generated noise, the experiment was repeated 10 times and mean value was taken as final result. Test results of type I noise and type II noise are shown in Table V.

TABLE V
EVALUATION RESULTS OF THE MODEL UNDER NOISE

Noise type	Noise level	MAE	RMSE	MAPE	R2
–	–	0.0577	0.0743	0.0331	0.7662
Type I	±3%	0.0622	0.0829	0.0350	0.7140
	±4%	0.0644	0.0858	0.0362	0.6935
	±5%	0.0667	0.0887	0.0376	0.6717
Type II	±3%	0.0571	0.0812	0.0321	0.7273
	±4%	0.0591	0.0839	0.0333	0.7095
	±5%	0.0602	0.0855	0.0339	0.6988

As can be seen from Table V, MAE, RMSE, MAPE and R2 only decrease by 0.009, 0.0144, 0.0045 and 0.0945 even if K of type I noise is ±5%. Under type II noise, performance of the model is relatively robust due to its randomness and small probability of addition. Under two types of noise, PLE shows good anti-noise performance and has high practical application value. This is due to the fact that the parameters of the model are shared by multiple tasks, which reduces over-fitting of the model to a specific task, thus enhancing anti-noise ability of the model.

V. CONCLUSION

This paper proposes an SSE method based on MTL and MRSCR, and tests it on the CEPRI-FS-102 bus system. Conclusions are as follows:

1) Analyze evaluation process of SSE method based on MRSCR, and determine the process to be enhanced is the calculation of CSCR. On this basis, the simulation process of CSCR dataset is proposed, and a corresponding simulation program is developed.

2) According to characteristics of the prediction task, MTL mode is used to complete synchronous prediction of CSCR of multiple RESs connection points, and training cost and evaluation time are greatly reduced. The used PLE can extract shared information and difference information between tasks, and has good predictive performance and anti-noise performance.

3) Compared with empirical value, the proposed method can accurately characterize critical state of the system and control evaluation error within a minimal range. Compared with theoretical values and off-line simulation, this method takes real-time measurement as the input characteristic and reduces evaluation time to millisecond level. Different from traditional research paradigms, this approach uses DL model to improve accuracy and practicability of the mechanism approach, and achieves a deep integration of data-driven and knowledge-driven.

This paper preliminarily explores the possibility of applying DL technology to SSE field in a way of intelligent enhancement, and provides new ideas for related research in this field. Future work will further analyze other weaknesses of the mechanism approach and verify applicability of hybrid intelligence enhancement approach.

REFERENCES

[1] S. Rahman, S. Saha, S. N. Islam, M. T. Arif, M. Mosadeghy, M. E. Haque, and A. M. T. Oo, “Analysis of power grid voltage stability with high penetration of solar PV systems,” *IEEE Transactions on Industry Applications*, vol. 57, no. 3, pp. 2245–2257, May/Jun. 2021.

[2] V. Telukunta, J. Pradhan, A. Agrawal, M. Singh, and S. G. Srivani, “Protection challenges under bulk penetration of renewable energy resources in power systems: a review,” *CSEE Journal of Power and Energy Systems*, vol. 3, no. 4, pp. 365–379, Dec. 2017.

[3] N. P. W. Strachan and D. Jovcic, “Stability of a variable-speed permanent magnet wind generator with weak AC grids,” *IEEE Transactions on Power Delivery*, vol. 25, no. 4, pp. 2779–2788, Oct. 2010.

[4] L. J. Cai and I. Erlich, “Doubly fed induction generator controller design for the stable operation in weak grids,” *IEEE Transactions on Sustainable Energy*, vol. 6, no. 3, pp. 1078–1084, Jul. 2015.

[5] H. J. Gu, R. F. Yan and T. Saha, “Review of system strength and inertia requirements for the national electricity market of Australia,” *CSEE Journal of Power and Energy Systems*, vol. 5, no. 3, pp. 295–305, Sep. 2019.

[6] F. Zhang, H. H. Xin, D. Wu, Z. Wang, and D. Q. Gan, “Assessing

strength of multi-infeed LCC-HVDC systems using generalized short-circuit ratio," *IEEE Transactions on Power Systems*, vol. 34, no. 1, pp. 467–480, Jan. 2019.
[7] H. Xiao, Y. Zhang, X. Z. Duan, and Y. H. Li, "Evaluating strength of hybrid multi-infeed HVDC systems for planning studies using hybrid multi-infeed interactive effective short-circuit ratio," *IEEE Transactions on Power Delivery*, vol. 36, no. 4, pp. 2129–2144, Aug. 2021.
[8] P. F. de Toledo, B. Bergdahl, and G. Asplund, "Multiple infeed short circuit ratio-aspects related to multiple HVDC into one AC network," in *2005 IEEE/PES Transmission & Distribution Conference & Exposition: Asia and Pacific*, 2005, pp. 1–6.
[9] H. D. Sun, S. Y. Xu, T. Xu, Q. Guo, J. B. He, B. Zhao, L. Yu, Y. Zhang, W. F. Li, Y. K. Zhou, Y. T. Zhang, and Y. Y. Zhu, "Definition and index of short circuit ratio for multiple renewable energy stations," *Proceedings of the CSEE*, vol. 41, no. 2, pp. 497–505, Jan. 2021.
[10] J. Schmall, S. H. Huang, Y. Li, J. Billo, J. Conto, and Y. Zhang, "Voltage stability of large-scale wind plants integrated in weak networks: an ERCOT case study," in *2015 IEEE Power & Energy Society General Meeting*, 2015, pp. 1–5.
[11] GE Energy Consulting. (2014, Oct.). Minnesota renewable energy integration and transmission study: final report [Online]. Available: https://mn.gov/commerce-stat/pdfs/mrits-report-2014.pdf.
[12] J. Bech, "Connection of wind farms to weak AC networks," CIGRE, Paris, Dec. 2016.
[13] L. Yu, H. D. Sun, S. Y. Xu, B. Zhao, and J. Zhang, "A critical system strength evaluation of a power system with high penetration of renewable energy generations," *CSEE Journal of Power and Energy Systems*, vol. 8, no. 3, pp. 710–720, May 2022.
[14] D. Wu, G. G. Li, M. Javadi, A. M. Malyscheff, M. G. Hong, and J. N. Jiang, "Assessing impact of renewable energy integration on system strength using site-dependent short circuit ratio," *IEEE Transactions on Sustainable Energy*, vol. 9, no. 3, pp. 1072–1080, Jul. 2018.
[15] W. Dong, H. H. Xin, D. Wu, and L. B. Huang, "Small signal stability analysis of multi-infeed power electronic systems based on grid strength assessment," *IEEE Transactions on Power Systems*, vol. 34, no. 2, pp. 1393–1403, Mar. 2019.
[16] J. J. Q. Yu, D. J. Hill, A. Y. S. Lam, J. T. Gu, and V. O. K. Li, "Intelligent time-adaptive transient stability assessment system," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 1049–1058, Jan. 2018.
[17] C. Ren, Y. Xu, J. H. 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 Journal of Power and Energy Systems*, vol. 8, no. 1, pp. 76–85, Jan. 2022.
[18] J. Xie and W. Sun, "Distributional deep reinforcement learning-based emergency frequency control," *IEEE Transactions on Power Systems*, vol. 37, no. 4, pp. 2720–2730, Jul. 2022.
[19] J. Li, S. Chen, X. Wang and T. Pu, "Load shedding control strategyin power grid emergency state based on deep reinforcement learning," *CSEE Journal of Power and Energy Systems*, vol. 8, no. 4, pp. 1175–1182, Jul. 2022
[20] T. J. Wang and Y. Tang, "Transient stability preventive control basedon graph convolution neural network and transfer deep reinforcementlearning," *CSEE Journal of Power and Energy Systems*, doi: 10.17775/CSEEJPES.2022.05030.
[21] D. H. A. Lee and G. Andersson, "An equivalent single-infeed model of multi-infeed HVDC systems for voltage and power stability analysis," *IEEE Transactions on Power Delivery*, vol. 31, no. 1, pp. 303–312, Feb. 2016.
[22] H. Y. Tang, J. N. Liu, M. Zhao, and X. D. Gong, "Progressive layered extraction (PLE): a novel multi-task learning (MTL) model for personalized recommendations," in *Proceedings of the 14th ACM Conference on Recommender Systems*, 2020, pp. 269–278.
[23] J. Q. Ma, Z. Zhao, X. Y. Yi, J. L. Chen, L. C. Hong, and E. H. Chi, "Modeling task relationships in multi-task learning with multi-gate mixture-of-experts," in *2018 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2018, pp. 1930–1939.
[24] A. Gupta, G. Gurrala, and P. S. Sastry, "An online power system stability monitoring system using convolutional neural networks," *IEEE Transactions on Power Systems*, vol. 34, no. 2, pp. 864–872, Mar. 2019.

Baoluo Li received the M.S. degree from Northeast Electric Power University, Jilin, China in 2021. He is currently pursuing the Ph.D. degree with the School of Electrical Engineering, Shandong University, Jinan, China. His current research interests include power system security, stability analysis, and machine learning.

Shiyu Xu received received the B.S. degree from Yanshan University, Qinhuangdao, China, in 2005 and the Ph.D. degree in Mechanical Systems and Control from Peking University, Beijing, China, in 2010. She worked as a Postdoctoral fellow with the China Electric Power Research Institute, Beijing, China, till 2012. From 2007 to 2008, she was a Visiting Scholar of Polytechnic Institute, New York University, USA. Currently, she is a Senior Engineer at China Electric Power Research Institute, Beijing, China. Her research interests include power systems dynamic stability analysis and control with high penetration of renewable energies.

Huadong Sun received the B.Eng. and M.S. degrees in Electrical Engineering from Shandong University, Jinan, China, in 1999 and 2002, respectively. He received the Ph.D. degree in Electrical Engineering from China Electric Power Research Institute (CEPRI) in 2005. He is currently the Vice President and a Professional Engineer of CEPRI. His research interests include power system analysis and control.

Zonghan Li received the B.Eng. and M.S. degrees in Electrical Engineering from Northeast Electric Power University, Jilin, China, in 2015 and 2018, respectively. He received the Ph.D. degree in Electrical Engineering from China Electric Power Research Institute, Beijing, China in 2021. He is currently working as a Postdoc in the Power System Department of China Electric Power Research Institute, Beijing, China. His main research interest includes the stability of power system.

Lin Yu received the B.Eng. and M.S. degrees in Electrical Engineering from Shandong University, Jinan, China, in 2014 and 2017, respectively. She received the Ph.D. degree in Electrical Engineering from China Electric Power Research Institute, Beijing, China in 2022. She is currently working as a Postdoc in the Power System Department of China Electric Power Research Institute, Beijing, China. Her main research interest includes the stability of power system with renewable energy.