Frontiers
2025, Vol. 5, No. 1, pp. 30-41
https://doi.org/10.11648/j.frontiers.20250501.12

Research Article

Data-Driven Frequency Security Assessment Based on Generative Adversarial Networks and Metric Learning

Li Huarui*, Zhu Xinyao, Jia Yongyong, Li Zheng, Wang Xiaobo

State Grid Jiangsu Electric Power Company Limited, Research Institute, Nanjing, China

Abstract

With construction of large-capacity direct current transmission projects and large-scale integration of renewable energy, frequency security of the power system is facing severe challenges. For fast and accurate online assessment of frequency security, a data-driven frequency security assessment model based on Generative Adversarial Network (GAN) and Metric Learning (ML) is proposed in this paper. Firstly, the key frequency security indicators are selected as the outputs of the model, and the input feature set is constructed. Then, distribution information of historical operation scenarios is learned through Wasserstein Generative Adversarial Network (WGAN), in order to generate operation scenarios covering typical operation modes for training sample set establishment. The generated operation scenarios are adjusted based on rejection sampling and resampling techniques, in order to increase the density of training samples near key scenes. Finally, considering inapplicability of a single assessment model for frequency security assessment in power systems with complicated changes of operation conditions, a combined assessment model for frequency security assessment composed of multiple sub-models is constructed based on Metric Learning for Kernel Regression (MLKR). The original distance metric is adjusted with metric learning techniques to make samples with similar frequency dynamics close. Then the samples with similar frequency dynamics are clustered into the same cluster, and the corresponding sub-model is established. A simplified Shandong power system example is used to verify the effectiveness of the proposed method.

Keywords

Frequency Security, Machine Learning, Data-Driven, Generative Adversarial Network, Metric Learning

1. Introduction

Driven by the development of High Voltage Direct Current(HVDC) transmission and renewable energy generation, the equivalent inertia and rotating reserve capacity of grids decreases continuously. Meanwhile, risk of accidents with large power shortage increases [1]. Frequency security of modern power system faces great challenges after accidents with large power shortages such as DC blocking [2]. There is an urgent need to carry out research on frequency security assessment methods for enhancing frequency security stability of power system [3].

Full-time domain simulation analyses frequency dynamics based on detailed models. Literature [4] takes into account impacts of boiler thermal dynamic process on unit output power to improve assessment accurac. Literature [5] establishes a detailed model describing frequency dynamics of large thermal power units for analysing frequency dynamic

*Corresponding author: lihuarui_sdu@163.com (Li Huarui)

Received: 31 October 2024; **Accepted:** 4 January 2025; **Published:** 24 January 2025

process after unloading. The time-domain simulation method takes into account the detailed model of power system, which can comprehensively reflect frequency dynamics [6]. However, in this process, algebraic differential equations need to be solved to accurately simulate the frequency response process after active perturbation, which is quite time-consuming. Therefore, full-time domain simulation is unsuitable for online frequency security assessment [7].

In order to speed up the assessment process, some simplified models are applied for frequency security assessment. Literature [8] proposes the Average System Frequency (ASF) model that ignores voltage dynamics and network effects. Further simplification based on the ASF model leads to the System Frequency Response (SFR) model [9]. Based on the SFR model, the analytical solution of the system frequency response can be obtained directly without stepwise integration. Literature [10] proposes a new method for analysing the frequency response of power systems based on DC power flow calculation by ignoring the influence of reactive-voltage variations on active-frequency dynamics. The simplified models above ignore network influence and coupling relationship between frequency dynamics and voltage dynamics, which leads to poor accuracy for assessing large-scale frequency deviation after severe power disturbance.

In recent years, machine learning methods have been widely applied in the field of online frequency security assessment. Once offline training completed, the assessment model can be used for frequency security online assessment without solving algebraic differential equations. In literature [11] and [12], Support Vector Regression (SVR) is applied to assess the minimum frequency after accidents. In literature [13] and [14], Extreme Learning Machine (ELM) are applied to assess the minimum frequency after accidents. In literature [15], the frequency assessment method based on Random Forest (RF) is proposed.

Once offline training completed, the machine learning model can be used for online frequency security assessment without solving algebraic differential equations. Accuracy of the model depends on adequacy of training samples and reasonableness of the model structure. In most studies, operation states of power systems are changed randomly within a certain range for generating training samples. In this way, generalisation ability of the frequency security assessment model can't be guaranteed. For different operation scenarios, frequency dynamics after serious power disturbance may differ greatly. Large assessment errors may arise with a single machine learning model.

Considering the above problems, a data-driven frequency security online assessment method based on Generative Adversarial Network(GAN) and Metric Learning(ML) is proposed. Firstly, the corresponding frequency security indicators are selected and input features are constructed; Then, distribution information of historical operation scenarios is learned through Wasserstein Generative Adversarial Network(WGAN), in order to generate that operation scenarios cover typical operation modes of power systems; Finally, considering inapplicability of a single assessment model, a combined assessment model for frequency security assessment composed of multiple sub-models is constructed based on Metric Learning for Kernel Regression (MLKR). Based on a simplified example of Shandong Power Grid, it is verified that the data-driven frequency security assessment model based on GAN and ML proposed in this paper can quickly and accurately assess the frequency security after severe power disturbances such like DC blocking under different operating scenarios.

2. Frequency Security Indicator Selection and Input Feature Construction

2.1. Frequency Security Indicator Selection

The frequency assessment model developed in this paper focuses on transient frequency security of the power system after a sudden active power disturbance. The maximum transient frequency deviation $\Delta f_{\max}$, the maximum frequency rate of change $f_{\max}^{\mathrm{ROCOF}}$ and the quasi-steady-state frequency deviation Δf_{ss} are important indicators reflecting frequency security of power systems after power disturbance [16, 17]. $\Delta f_{\max}$ is related to the inertia response process of units after the occurrence of power disturbance events, which determines action of relevant system protections; $f_{\max}^{\mathrm{ROCOF}}$ is related to the amount of power disturbance caused by power disturbance events and inertia of the system; Δf_{ss} can be used to judge whether the steady frequency satisfies the corresponding requirements. In this paper, the data-driven frequency security assessment model is constructed with $\Delta f_{\max}$, $f_{\max}^{\mathrm{ROCOF}}$ and Δf_{ss} as frequency security indicators.

2.2. Initial Input Feature Construction

Both the pre-disturbance power flow features reflecting the information of the system operation mode and the system dynamic features reflecting the frequency dynamic response process after the power disturbance can be used as input features of the frequency safety assessment model. However, the acquisition of post-disturbance dynamic information for online assessment of system frequency safety relies on time-domain simulation, which will affect the speed of model assessment. In order to meet the requirement of rapidity of assessment, this paper selects the pre-disturbance power flow features to construct the input features of the frequency safety assessment model, so as to avoid the dependence on the time domain simulation and ensure the rapidity of assessment.

The frequency of power system is determined by the the inertia equation of units:

$$\frac{d\Delta f(t)}{dt} = \frac{1}{2H}\left(\sum_{i=1}^{N_G} \Delta P_G^i - \Delta P_d - D\Delta f(t)\right) \quad (1)$$

Where $\Delta f(t)$ represents the frequency deviatio, H represents the inertia time constant, ΔP_G^i represents the active power adjustment associated with the dynamic response of the prime mover-governor system, ΔP_d rrepresents the power disturbance, D represents damping of the system, and N_G represents the number of units.

For the three key indicators of frequency dynamics after disturbance selected in this paper, Δf_{max}, f_{max}^{ROCOF} and Δf_{ss}, they are closely related to the amount of disturbance and the system inertia level, and the larger the amount of disturbance and the lower the system inertia level are, the faster the change is; for Δf_{max}, the magnitude of its value is closely related to inertia level of the system, the reserve capacity, and regulation rate of units; and Δf_{ss} is mainly affected by factors such as the size of the power disturbance caused by the corresponding event, and the level of the reserve capacity of the power system itself. The input characteristics should be as inclusive as possible of the operating characteristics of the power system. Based on the above analysis, this paper selects the conventional unit output, reserve capacity, unit capacity, regulated power pwer unit, inertia time constant, output of the renewable energy units, system damping coefficient, load size, DC power, and disturbance amount as the original input features before the disturbance.

3. Training Sample Generation Based on Generative Adversarial Network

The frequency of serious power disturbances such as DC blocking is extremely low in the actual operation of the power system, and it is difficult to collect enough scenario information from the historical operation data of the power system to generate training samples. It is necessary to construct a set of power system operation scenarios in the offline stage to simulate the frequency dynamics after power disturbances in the time domain, so as to obtain the frequency security indexes such as the maximum frequency deviation of the system under the corresponding scenarios. And then training samples for the training of the frequency security assessment model are generated. Currently, the training sample generation method usually adopts the method of randomly varying the system operating states such as load level and new energy output within a certain range, so as to generate a batch of operating scenarios to construct the training sample set [18]. This training sample generation method is difficult to cover the actual operation of the power system, and cannot guarantee the ability of the frequency security assessment model to generalise to the future scenarios to be assessed. And it will generate a large number of training samples that are far away from the actual operation of the system, resulting in low efficiency of model training.

Generative Adversarial Network (GAN) [19], as a data-driven sample distribution learning method, can learn the sample distribution law without assuming the sample distribution function. In this paper, based on the information of the historical operation mode of the power system, we use the improved Wasserstein Distance Metric-based Generative Adversarial Network (Wasserstein Generative Adversarial Network, WGAN) to learn the distribution law of the actual operation scenario of the system. Then a batch of scenarios in line with the actual operation of the system is generated for time-domain simulation to obtain the frequency security index after the predicted power perturbation event, which is used to construct the training sample set.

3.1. Operational Scene Generation Based on Conventional GAN Networks

Traditional GAN is an unsupervised learning method with a generator and a discriminator [19]. For the GAN used for power system operation scene generation, its generator simulates and generates the operation scene S that is as close as possible to the actual operation law of the power system by learning the historical operation scene distribution of the power system. The discriminator is used to judge whether the generated operation scene is close to the real scene distribution law of the power system. $P_{his}(S)$ represents the realistic distribution of historical operating scenarios for the power system. Given a noise distribution $P_{noise}(S)$, the goal of GAN is to map the sampled data obtained via $P_{noise}(S)$ through the game process of generator and discriminator to make it as close as possible to $P_{his}(S)$.

Operation characteristics of the power system include network topology characteristics, generation load pattern characteristics, and so on. The frequency security assessment model proposed in this paper is used to assess the frequency security indexes under a large number of different source-load scenarios corresponding to the operation modes, so the changes in the generation and load pattern information are mainly taken into account when constructing the operation scenarios, which are characterised by the power flow feature vectors composed of wind power, photovoltaic power, load and conventional thermal power unit power, and so on.

For the generator $G(\bullet;W^G)$, the input is the sampled noise data S_{noise} obtained via $P_{noise}(S)$ and the output is the generated data samples $G(S_{noise};W^G)$ that simulate the distribution pattern of the real historical operating scenarios of the power system. The objective of the generator is to make the distribution of $G(S_{noise};W^G)$, which is represented by $P(G(S_{noise};W^G))$, as close as possible to the distribution of

the real historical operating scenarios of the power system $P_{\text{his}}(\text{S})$, and its loss function L_{G} can be expressed as follows:

$$L_{\text{G}} = -E_{\text{S}_{\text{noise}} \sim P_{\text{noise}}(\text{S})}(D(G(\text{S}_{\text{noise}}; W^{\text{G}}))) \tag{2}$$

where $E()$ represents the expectation and $D()$ represents output of the discriminator network.

For the discriminator, its input is the real historical operation scene of the power system or the simulated historical operation scene generated by the generator, and the output is the probability value of judging that the corresponding operation scene is the real historical operation scene, i.e., the goal of the discriminator network is to correctly discriminate the source of the input operation scene as much as possible. Its loss function is denoted as:

$$L_{\text{D}} = -E_{\text{S}_{\text{his}} \sim P_{\text{his}}(\text{S})}(D(\text{S}_{\text{his}})) + E_{\text{S}_{\text{noise}} \sim P_{\text{noise}}(\text{S})}(D(G(\text{S}_{\text{noise}}))) \tag{3}$$

In order to model the game between the generator and the discriminator so that they can be trained simultaneously, the game in Eq. (4) is modelled.

$$\begin{aligned} &\min_{G} \max_{D} V(G,D) = \\ &E_{\text{S}_{\text{his}} \sim P_{\text{his}}(\text{S})}(D(\text{S}_{\text{his}})) - E_{\text{S}_{\text{noise}} \sim P_{\text{noise}}(\text{S})}(D(G(\text{S}_{\text{noise}}))) \end{aligned} \tag{4}$$

Through the game process between the generator and the discriminator, the generator that simulates the distribution law of the real historical operation scenario of the power system is finally obtained.

3.2. Operational Scenario Generation Based on Improved WGAN Network

The traditional GAN model has problems such as gradient vanishing and pattern collapsing, this paper further introduces an improved WGAN network model based on Wasserstein distance metric [20] to learn the distribution law of power system operation scene. The Wasserstein distance between and is calculated as shown in equation (5):

$$\begin{aligned} &W(P(G(\text{S}_{\text{noise}}; W^{G})), P_{\text{his}}(\text{S})) \\ &= \inf_{\gamma \sim \Pi(P(G(\text{S}_{\text{noise}}; W^{G})), P_{\text{his}}(\text{S}))} E_{(u,v)\sim\gamma}[\|u - v\|] \end{aligned} \tag{5}$$

where $\Pi(P(G(\text{S}_{\text{noise}}; W^{G})), P_{\text{his}}(\text{S}))$ represents the set of joint probability distributions γ with $P(G(\text{S}_{\text{noise}}; W^{G}))$ and $P_{\text{his}}(\text{S})$ as marginal distributions, and inf represents the maximum lower bound.

There are difficulties in calculating the Wasserstein distance directly. In this paper, the Kantorovich-Rubinstein dyadic form is used to measure the Wasserstein distance, as shown in equation (6):

$$\begin{aligned} &W(P(G(\text{S}_{\text{noise}}; W^{G})), P_{\text{his}}(\text{S})) = \\ &\sup_{\|D\|_L \le 1} E_{\text{S}\sim P_{\text{his}}(\text{S})}[D(\text{S})] - E_{\text{S}\sim P_{\text{noise}}(\text{S})}[D(G(\text{S}))] \end{aligned} \tag{6}$$

Where sup denotes the upper bound, i.e., the minimum upper bound, $\|D\|_L \le 1$ indicates that the discriminator network needs to satisfy the 1-Lipschitz conditional restriction, i.e., constraining the variation of the network parameters within a certain range during the training time, in order to ensure the normal gradient optimisation.

3.3. Generative Scene Adjustment Based on Rejection Sampling and Resampling Techniques

Based on the simulated power system operation scenarios generated by the improved WGAN network. Detailed time-domain simulations of frequency dynamics after severe power disturbances, such as DC blocking, are performed to obtain the corresponding frequency security indicators, such as the maximum frequency offset. So that the training sample set for the frequency security assessment model can be constructed. Some operating scenarios to be evaluated may lack training samples with similar operation modes. It's difficult to learn the complex mapping relationship between frequency safety indicators and powe flow characteristics under the corresponding operation modes based on sparse training samples. This makes the frequency safety assessment model has a large error in assessing the frequency safety of these scenarios. Improving the density of training samples near these operating scenarios is the key to reduce the frequency safety assessment error. In this paper, we define these operational scenarios with large assessment errors as key scenarios, and increase the density of training samples near the key scenarios through repeated rejection sampling [21] and resampling [22] to improve the accuracy of frequency safety assessment.

For the key scenarios $\text{S}_{\text{key},i}$, $\text{S}_{\text{key},j}$, and $\text{S}_{\text{key},k}$, the evaluation errors of the frequency safety indexes are e_i, e_j and e_k, $e_i \le e_j \le e_k$. According to the principle that the newly generated scenarios S_{new} should be as close as possible to the key samples with large evaluation errors, there should be:

$$d(\text{S}_{\text{new}}, \text{S}_{\text{key},i}) \ge d(\text{S}_{\text{new}}, \text{S}_{\text{key},j}) \ge d(\text{S}_{\text{new}}, \text{S}_{\text{key},k}) \tag{7}$$

Where $d()$ represents the distance between two scenes.

The sequence $\{ e_l, 1 \le l \le N_{\text{key}} \}$ and the sequence $\{ d(\text{S}_{\text{new}}, \text{S}_{\text{key},m}), 1 \le m \le N_{\text{key}} \}$ should have a high negative correlation to ensure that Eq. (7) holds, so as to generate training samples as close as possible to the key scenarios. In this paper,

Spearman correlation coefficient is introduced to measure the correlation of two sequences, as shown in equation (8):

$$\rho(\{e_l,1\le l\le N_{\text{key}}\},\{d(\text{S}_{\text{new}},\text{S}_{\text{key},m}),1\le m\le N_{\text{key}}\})=\frac{\sum_{n=1}^{N_{\text{key}}}[R(e_n)-\bar{R(e)}][R(d(\text{S}_{\text{new}},\text{S}_{\text{key},n}))-\bar{R(d)}]}{\sqrt{\sum_{n=1}^{N_{\text{key}}}[R(e_n)-\bar{R(e)}]^2\sum_{n=1}^{N_{\text{key}}}[R(d(\text{S}_{\text{new}},\text{S}_{\text{key},n}))-\bar{R(d)}]^2}} \quad (8)$$

$$\bar{R(e)}=\frac{1}{N_{\text{key}}}\sum_{n=1}^{N_{\text{key}}}R(e_n)$$

$$\bar{R(d)}=\frac{1}{N_{\text{key}}}\sum_{n=1}^{N_{\text{key}}}R(d(\text{S}_{\text{new}},\text{S}_{\text{key},n}))$$

Where $R(e_n)$ and $R(d(\text{S}_{\text{new}},\text{S}_{\text{key,n}}))$ represent the rank of the sequence $\{e_l,1\le l\le N_{\text{key}}\}$ and the corresponding element in the sequence $\{d(\text{S}_{\text{new}},\text{S}_{\text{key},m}),1\le m\le N_{\text{key}}\}$, respectively.

Setting the threshold ρ_{th}, the rejection sampling function shown in equation (9) is constructed to screen out the training samples away from the critical scene.

$$h(\text{S}_{\text{new}})=\begin{cases}1, & \rho(\{e_l,1\le l\le N_{\text{key}}\},\\ & \{d(\text{S}_{\text{new}},\text{S}_{\text{key},m}),1\le m\le N_{\text{key}}\})<\rho_{\text{th}}\\ 0, & \rho(\{e_l,1\le l\le N_{\text{key}}\},\\ & \{d(\text{S}_{\text{new}},\text{S}_{\text{key},m}),1\le m\le N_{\text{key}}\})\ge\rho_{\text{th}}\end{cases} \quad (9)$$

Each time a new scenario is added based on the generator $G(\bullet;W^{\text{G}})$ to simulate the distribution pattern of historical actual operating scenarios of the power system. New scenes are randomly divided to generate training samples and test samples, respectively. The scenarios corresponding to the test samples with test errors larger than a set threshold are defined as critical scenarios. Based on the WGAN network generator and the rejection sampling function in Eq. (9), the density of training samples near the key scenes can be increased by rejection sampling and resampling.

4. Frequency Safety Combined Assessment Model Structure and Learning Process

4.1. Structure of the Frequency Safety Combined Assessment Model

The frequency safety machine learning assessment model learns the functional relationship $y=F(x)$ between the frequency safety index y and the input features x based on historical training samples, so that the frequency safety index of the scenario to be assessed can be obtained directly based on $y=F(x)$ without solving the set of differential algebraic equations. Among them, y can be taken as $\Delta f_{\max}$, $f_{\max}^{\text{ROCOF}}$ or Δf_{ss}. x is the input feature of the model formed after the feature dimensionality reduction of the original current features. The power system inertia level, spinning reserve level, blocking DC, etc. may differ greatly under different operating scenarios, resulting in significant differences in the system frequency dynamic process after perturbation under different scenarios. In this case, the functional expression $y=F(x)$ describing the relationship between the frequency safety index and the input features under different scenarios may have significant differences, which makes it difficult to learn $y=F(x)$ using a single machine learning model. In this paper, a combined frequency safety assessment model consisting of several sub-models is proposed for assessing the frequency safety of the system after severe disturbances such as DC blocking.

In the process of establishing the frequency safety combination assessment model, the samples are clustered based on the principle of grouping the samples with similar mapping relationships between the frequency safety indexes and the input features into the same cluster as far as possible. and then the corresponding sub-models are established for each cluster. Traditional clustering algorithms are usually based on the distance between samples, according to the principle that the distance between samples in the same cluster is as small as possible while the distance between samples in different clusters is as large as possible for sample clustering. However, it is difficult to measure the similarity of the mapping relationship between the frequency safety index and the input features by the commonly used distance. In this paper, Metric Learning (ML) method is introduced to adjust the original distance metric, so that the adjusted distance reflects the similarity between the frequency safety indexes of different samples and the function mapping relationship between the input features, and then the traditional distance-based clustering algorithm is used to divide the similar samples into the same cluster.

The process of frequency safety combination assessment model building can be divided into three steps: firstly, the original distance metric is adjusted with the goal of reflecting the similarity of the function mapping relationship between different samples of frequency safety indexes and input features, i.e., the metric learning process; then, the similar samples are classified into the same clusters; and finally, the sub-models of frequency safety assessment are built for each cluster respectively. The whole process is shown in Figure 1, where the points of the same shape indicate the function mapping relationship similar samples.

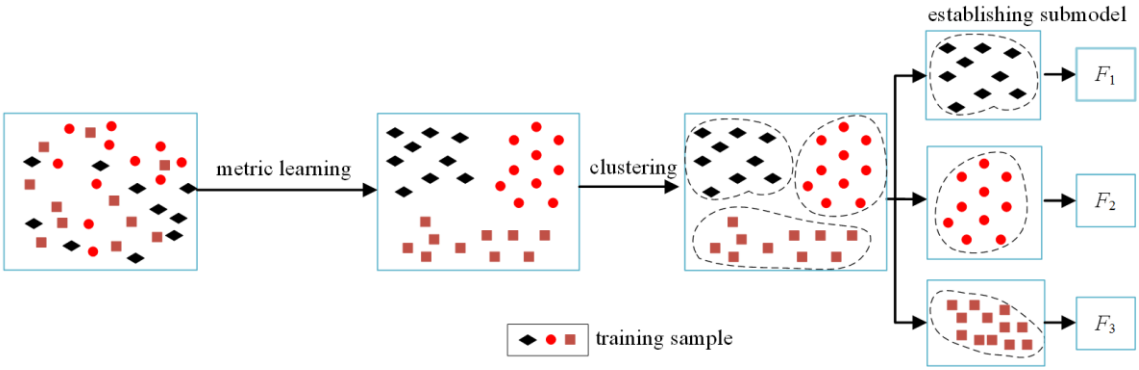


Figure 1. *The establishment process of the combined assessment model for frequency security assessment.*

4.2. Sample Distance Adjustment Based on MLKR Algorithm

Metric learning adjusts the original distance metric, and the adjusted distance metric reflects the similarity of the function mapping relationship between different samples' frequency safety indicators and the input features. The more similar the samples are, the closer they are to each other. Thus, the samples can be divided into clusters according to the adjusted distance metric, and the similar samples can be divided into the same cluster.

In this paper, Mahalanobis Distance [23] is introduced to measure the distance between different samples, and the Mahalanobis Distance is calculated as shown in Equation (10):

$$d_{\text{mah}}(\mathrm{x}_i,\mathrm{x}_j)=\sqrt{(\mathrm{x}_i-\mathrm{x}_j)^{\mathrm{T}}\mathrm{M}(\mathrm{x}_i-\mathrm{x}_j)}=\sqrt{\left\|\mathrm{x}_i-\mathrm{x}_j\right\|_{\mathrm{M}}^2} \quad (10)$$

Where M is a semi-positive definite transformation matrix that can be tuned by metric learning, i.e., a martensian matrix.

The metric learning process is the matrix M adjustment process. In the metric learning process it is firstly necessary to choose the $F()$ form, i.e., to choose the appropriate machine learning method embedded in the metric learning process. Lazy Learning (LL) does not require an explicit training process [24] and is more suitable to be embedded in the metric learning process compared to Eager Learning (EL). In this paper, the classical lazy learning algorithm Kernel Regression (KR) algorithm is chosen to be embedded within the metric learning. The metric learning process for that is Metric Learning for Kernel Regression (MLKR) [24].

In this paper, the Gaussian kernel function [24] is chosen as the KR kernel function. Its radius is set to 1. The evaluation value of the frequency security index based on the KR method $\hat{y}_i$ is shown in Equation (11):

$$\hat{y}_i=\frac{\sum_{j=1}^{m} y_j e^{-d_{\text{mah}}(\mathrm{x}_i,\mathrm{x}_j)}}{\sum_{j=1}^{m} e^{-d_{\text{mah}}(\mathrm{x}_i,\mathrm{x}_j)}} \quad (11)$$

where y_j is the true value of the frequency safety index for the jth historical training sample and m is the number of training samples.

The loss function of the MLKR learning process is shown in Eq. (12), and the MLKR learning process is the optimization problem solving process of adjusting the matrix to minimize the loss function in Eq. (12).

$$\min L=\sum_{i=1}^{m}\left(y_i-\hat{y}_i\right)^2 \quad (12)$$

The semi-positive definite matrix M is decomposed into the form as in equation (13):

$$\mathrm{M}=\mathrm{A}^{\mathrm{T}}\mathrm{A} \quad (13)$$

The Mahalanobis distance is redefined into the form shown in equation (14), and the MLKR learning process, i.e., the process of solving the matrix with the objective of minimising the loss function in equation (12).

$$\begin{aligned} d_{\text{mah}}(\mathrm{x}_i,\mathrm{x}_j)&=\sqrt{(\mathrm{x}_i-\mathrm{x}_j)^{\mathrm{T}}\mathrm{A}^{\mathrm{T}}\mathrm{A}(\mathrm{x}_i-\mathrm{x}_j)} \\ &=\sqrt{\left\|\mathrm{A}(\mathrm{x}_i-\mathrm{x}_j)\right\|^2} \end{aligned} \quad (14)$$

In this paper, the Batch Gradient Descent (BGD) method [25] is used to solve A. The correction amount ΔA for each iteration of A during MLKR learning is calculated as shown in equations (15) and (16):

$$\frac{\partial L}{\partial \mathrm{A}} = 4\mathrm{A}\sum_{i}(\hat{y}_i - y_i)\sum_{j}(\hat{y}_j - y_j)k_{ij}\mathrm{x}_{ij}\mathrm{x}_{ij}^{\mathrm{T}} \tag{15}$$

$$\Delta \mathrm{A} = -\varepsilon\frac{\partial L}{\partial \mathrm{A}} \tag{16}$$

where ε represents the step size.

4.3. Sample Clustering and Sub-model Construction

After MLKR learning to adjust the distance between samples, the samples with similar frequency safety metrics and function mapping relationships between input features are close. Further clustering algorithm is needed to classify the similar samples into the same cluster, so as to build the frequency safety sub-evaluation model for each cluster. In this paper, we use the classical fuzzy k-means clustering [26] to divide the samples into clusters. In fuzzy k-means clustering, the initial cluster centroids (j=1,2,.....,K) are randomly selected from the samples, where K is the number of clusters. Then the affiliation degree u(x_i ,q) of the sample x_i to the qth sample cluster is calculated as shown in equation (17):

$$u(\mathrm{x}_i,q)=\left(\sum_{j=1}^{K}\left(\frac{d_{\mathrm{mah}}(\mathrm{x}_i,\mathrm{c}_q)}{d_{\mathrm{mah}}(\mathrm{x}_i,\mathrm{c}_j)}\right)^{\frac{1}{\beta-1}}\right)^{-1} \tag{17}$$

Where β is the set affiliation parameter.

Cluster centers are updated using the average of the qth cluster sample affiliation:

$$\mathrm{c}_q = \frac{\sum_{i=1}^{N_q}(u(\mathrm{x}_i,q))^{\beta}\mathrm{x}_i}{\sum_{i=1}^{N_q}(u(\mathrm{x}_i,q))^{\beta}} \tag{18}$$

Where N_q is the number of samples in the qth sample cluster.

The iterative process of splitting clusters is repeated until the centre of the sample clusters does not change any more, and the final cluster centre can be obtained. The affiliation of the scene to be evaluated to each cluster is defined as the confidence level of its belonging to each cluster, and the corresponding scene is assigned to the sample cluster with the highest confidence level.

Compared with lazy machine learning methods, eager machine learning methods construct explicit functional relationship expressions between frequency safety metrics, such as maximum frequency offset, and input features through offline learning, and the online evaluation of frequency safety is directly based on the constructed expressions, which makes the evaluation process rapid and does not need to carry out the calculation of distances between the scenarios to be evaluated and the training samples in lazy machine learning. Therefore, after completing the metric learning and sample clustering, this paper finally chooses the classical eager learning method SVR method [18] to construct the corresponding sub-evaluation model for each sample cluster. The expression between the frequency safety metrics and input features of the sub-evaluation model based on the SVR method is shown in Equation (19):

$$F(\mathrm{x}_i)=\mathrm{w}^{\mathrm{T}}\varphi(\mathrm{A}\mathrm{x}_i)+b \tag{19}$$

Where w is the coefficient vector, b is the bias, and φ is the mapping of functions corresponding to the kernel function, which is taken to be the Gaussian kernel function for the SVR model in this paper.

5. Case Study

5.1. Introduction of the Example System

The simplified Shandong power grid shown in Figure 2 is used as an example to verify the effectiveness of the frequency security machine learning assessment method proposed in this paper. The total load of the example grid is 58GW under the baseline operation mode, and there are three DC feeders, with a total DC feeder of 20GW, of which DC#1 and DC#2 lines have a capacity of 8GW, and DC#3 line has a capacity of 4GW. The No.1 wind farm, the No.2 wind farm, the No.3 wind farm, and the No.4 wind farm, are connected to nodes 170, 178, 187, and 322 in the example system respectively.

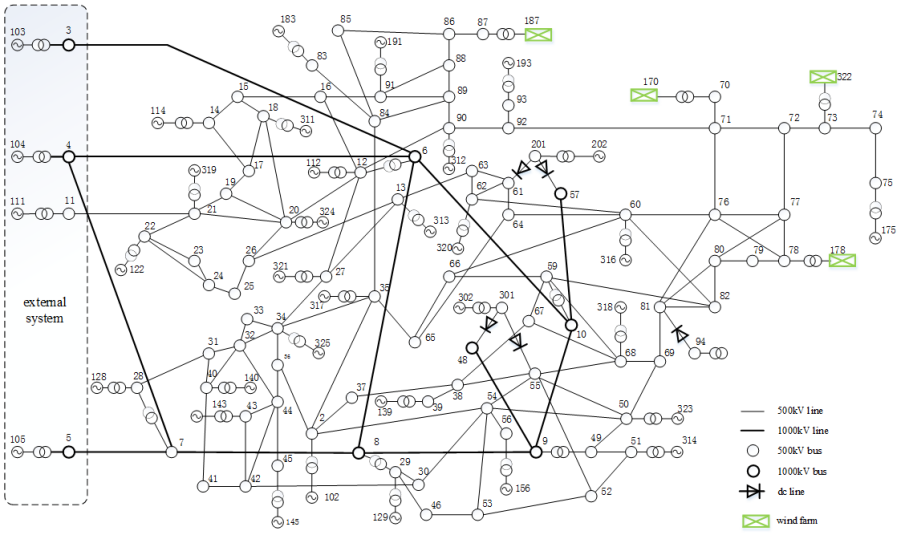


Figure 2. *The single line diagram of simplified Shandong grid.*

5.2. Sample Set Construction

Taking 15 min as the time interval, based on the wind power and load data under 2976 time sections of the system for 1 month, 2976 system operation scenarios are obtained by trend calculation with the goal of minimising network loss. Taking DC#2 blocking as the expected accident, 976 of the 2976 scenarios are taken as the historical operation scenarios. Time domain simulation is performed to obtain the frequency safety indexes under the expected accident to construct the original training sample set. Based on the remaining 2000 scenarios, time domain simulation is performed under the predicted accidents to construct the test sample set.

In this paper, the detailed learning process of the model with $\Delta f_{\max}$ as the output of the model assessment indicators is demonstrated. The learning process of the model with $f_{\max}^{\mathrm{ROCOF}}$ and Δf_{ss} as the assessment indicators is similar. In Section 4.6, the final assessment accuracy of the frequency security assessment model built based on the proposed method for each frequency security indicator is given uniformly.

5.3. Generation of New Training Samples

MLKR learning is performed based on 976 initial training samples, and the MLKR learning process converges with the assessment error converging to a larger value of about 0.137 Hz. New scenarios are generated using the method proposed in this paper to generate new training samples for model training to improve accuracy of the frequency safety assessment mode. A new set of operating scenarios is generated based on WGAN that conforms to the distribution pattern of historical scenarios, which is compared with the method in [22]. In order to visualise the comparison results, the scenarios are downscaled using the t-SNE method [27] into the form of two-dimensional feature representations of x_1 and x_2, as shown in Figure 3. From the figure, it can be seen that using the scene generation method via GAN proposed in this paper can generate scenes that follow the distribution law of the system's historica operation scenes.

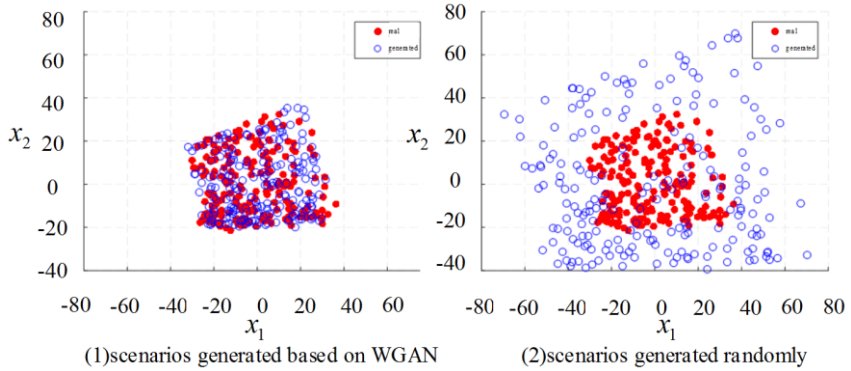


Figure 3. *Comparison of scenario distributions with different scene generation methods.*

The operation scenarios corresponding to the samples with evaluation errors greater than 0.1Hz based on the KR method are defined as key scenarios, and the newly generated scenarios are adjusted using the rejection sampling and resampling methods. By rejecting sampling and resampling, the density of training samples near the critical scenes is increased until the evaluation errors of all the scenes to be evaluated are less than 0.1Hz, and the training sample set after increasing the training samples is finally obtained with a total of 2507 training samples.

5.4. Learning Process of Combined Frequency Security Assessment Model

MLKR learning is performed based on the training sample set and the learning effect is tested using the test set. Before MLKR learning, the maximum training error of $\Delta f_{\max}$ is 0.187 Hz, and the maximum testing error is 0.169 Hz. after MLKR learning, the maximum training error is reduced to 0.0404 Hz, and the maximum testing error is reduced to 0.0468 Hz. With the process of MLKR learning, the changes of training error and testing error are shown in Figure 4, where $e_{\max}$ denotes the maximum evaluation error of all samples in each iteration. The evaluation error gradually converges as the number of MLKR learning iterations increases.

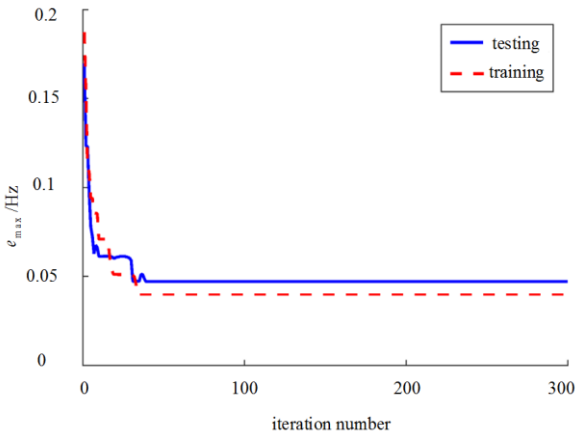
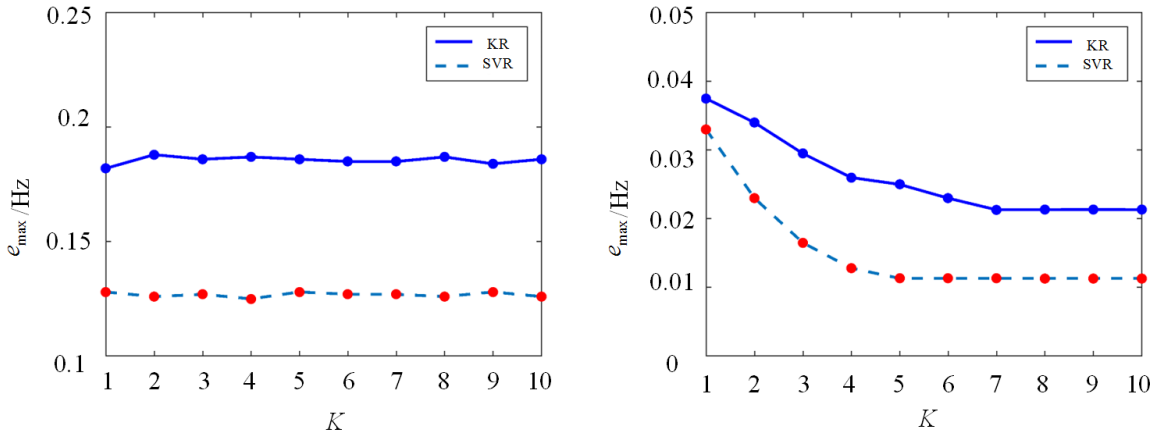


Figure 4. *The variation of the assessment error in the process of MLKR learning.*

After MLKR learning, the samples are clustered using fuzzy k-means clustering method and the corresponding evaluation sub-models for each cluster are built using the KR and SVR method, respectively. Before and after MLKR learning, with the increasement of the number of clusters *K*, the KR and SVR assessment errors are shown in Figure 5.



（1）before MLKR learning （2）after MLKR learning

Figure 5. *Error of the KR and SVR sub-models with different K before and after MLKR learning.*

Before MLKR learning, the frequency dynamics of samples close to each other are not necessarily similar, and clustering based on the original distance metric cannot ensure that samples with similar frequency safety metrics and input eigenfunction mapping relationships are grouped into the same cluster, and the frequency safety assessment error basically remains unchanged as the *K-value* increases. After learning by MLKR, the distance between samples is correlated with the dynamic similarity of frequency between samples, and the frequency safety indexes are close to the samples with similar mapping relationship with the input eigenfunction. With the increase of *K* value, the frequency dynamic similarity of samples in the same cluster increases, and the samples in each cluster are evaluated using KR and SVR, and the evaluation errors are all gradually reduced with the increase of *K* value, and the accuracy of the model evaluation is improved, as shown in Figure 5. Compared with the KR method, the SVR assessment model has better function fitting ability, and

therefore has higher frequency safety index assessment accuracy. With the increase of K value, when $K>6$, the model evaluation error is no longer significantly reduced, so this paper finally takes $K=6$.

5.5. Comparison with Other Methods

Different frequency security assessment models are compared with the combined assessment model proposed in this paper as shown in Figure 6, where ANN represents the assessment model based on Artificial Neural Network, FLNN represents the assessment model based on Functional Link Neural Network, v-SVR represents the assessment model based on v-Support Vector Regression (v-SVR), ELM represents the assessment model based on Extreme Learning Machine (ELM), ML represents the assessment model based on Metric Learning, and CAM represents the combined assessment model in this paper.

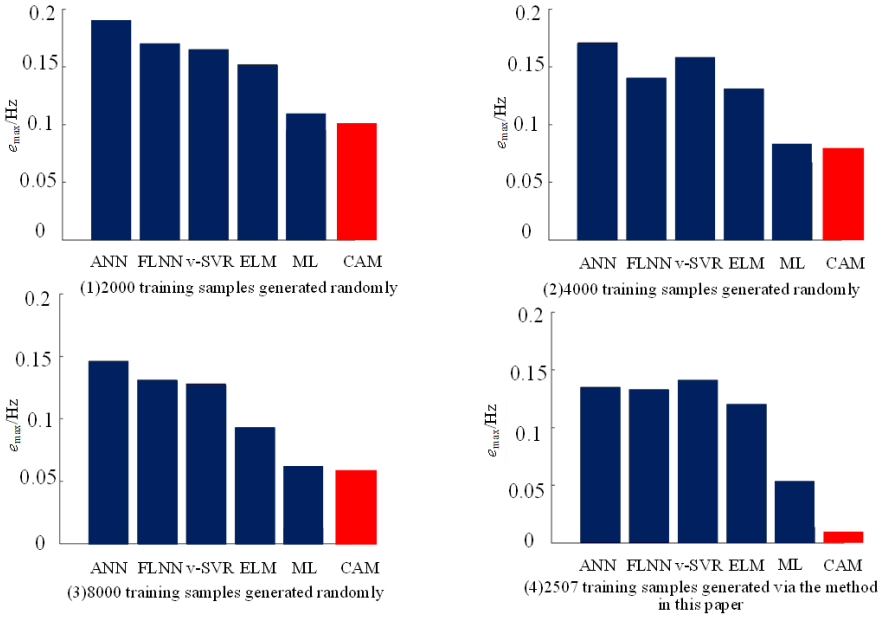


Figure 6. *The test errors of different models with different training sample sets.*

As shown in Figure 6, the combined frequency security assessment model proposed in this paper demonstrate better generalisation ability compared with other methods. The training samples obtained via the method in this paper can better cover the distribution of the actual operating scenarios. By using the rejection sampling and resampling techniques, the density of training samples near the key scenarios with large evaluation errors is increased, and the efficiency of model training is improved. Using the training sample generation method proposed in this paper, only 2507 training samples are needed to reduce the model testing error to 0.0113 Hz.

5.6. Accuracy of Assessment of Frequency Security Indicators

The accuracy of each frequency security indicators obtained based on the combined frequency security assessment model proposed in this paper is shown in Table 1. As shown in Table 1, the combination model for frequency security assessmentproposed in this paper demonstrate high accuracy for each frequency security indicator. Based on the method proposed in this paper, the frequency security indicators after severe power disturbance can be evaluated quickly and accurately.

Table 1. *The maximum and average values of the assessment errors of each frequency security indicator.*

frequency security indicator	Δf_{max}	f_{max}^{ROCOF}	Δf_{ss}
maximum of assessment errors	0.0113Hz	0.0098Hz•s^{-1}	0.0092Hz
average of assessment errors	0.0093Hz	0.0067Hz•s^{-1}	0.0071Hz

6. Conclusions

In this paper, a data-driven frequency security assessment model based on Generative Adversarial Network(GAN) and Metric Learning(ML) is proposed. In terms of frequency security indicatos selection and input feature construction,

key indicators reflecting frequency dynamic process after perturbation are selected, and pre-perturbation power flow features are chosen to construct the input features. In terms of generation of training samples, the method of generating training samples based on GAN is proposed, and sampling techniques are applied to increase sample density near key scenes. In terms of assessment model structure, the combined model for frequency security assessment is proposed, in order to fully take into account the large differences in frequency dynamics after perturbation events in different scenario. Through the final simplified Shandong power grid example, it is verified that with the proposed method, frequency security can be quickly and accurately assessed after severe power disturbances such like DC blocking.

Abbreviations

GAN:	Generative Adversarial Network
ML:	Metric Learning
WGAN:	Wasserstein Generative Adversarial Network
MLKR	Metric Learning for Kernel Regression
HVDC	High Voltage Direct Current
ASF	Average System Frequency
SFR	System Frequency Response
SVR	Support Vector Regression
ELM	Extreme Learning Machine
RF	Random Forest
LL	Lazy Learning
EL	Eager Learning
KR	Kernel Regression

Acknowledgments

The authors would like to thank the financial support by National Key Research and Development Plan of China (2024YFB2408900).

Author Contributions

Li Huarui: Conceptualization, Data curation, Formal Analysis, Methodology, Software

Li Zheng: Writing – original draft

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] HU Jiawei, WANG Tong, WANG Zengping. Collaborative emergency control strategy of system transient stability after DC blocking [J]. Power System Protection and Control, 2023, 51(4): 43-52. https://doi.org/10.19783/j.cnki.pspc.220630

[2] XIANG Yuwei, WANG Tong, LI Congcong, et al. Strategy of emergency generator tripping control for transient stability after a large capacity HVDC blocking fault [J]. Power System Protection and Control, 2021, 49(15): 84-92. https://doi.org/10.19783/j.cnki.pspc.201172

[3] HAN Zelei, JU Ping, QIN Chuan, et al. Review and prospect of research on frequency security of new power system [J]. Electric Power Automation Equipment, 2023, 43(09): 112-124. https://doi.org/10.16081/j.epae.202303007

[4] SHI Xiang, LIU Hongbo, XU Xingwei, et al. Study of impact of boiler's dynamic characteristics on dynamic frequency process in power system [J]. Electric Power Automation Equipment, 2007, 27(5): 69-72.

[5] WANG Yizhen, MA Shiying, WANG Qing, et al. Dynamic frequency response characteristics of large thermal power generation units [J]. Power System Technology, 2013, 37(1): 106-111. https://doi.org/10.13335/j.1000-3673.pst.2013.01.021

[6] BEVRANI H, GHOSH A, LEDWICH G. Renewable energy sources and frequency regulation: survey and new perspectives [J]. IET Renewable Power Generation, 2010, 4(5): 438-457. https://doi.org/10.1049/iet-rpg.2009.0049

[7] MOGHADDAM S Z. Generation and transmission expansion planning with high penetration of wind farms considering spatial distribution of wind speed [J]. International Journal of Electrical Power & Energy Systems, 2019, 106(3): 232-241. https://doi.org/10.1016/j.ijepes.2018.10.007

[8] CHAN M L, DUNLOP R D, SCHWEPPW F. Dynamic equivalents for average system frequency behavior following major disturbances [J]. IEEE Transactions on Power Apparatus and Systems, 1972, 91(4): 1637-1642.

[9] ANDERSON P M, MIRHEYDAR M. A low-order system frequency response model [J]. IEEE Transactions on Power Systems, 1990, 5(3): 720–729.

[10] LI Changgang, LIU Yutian, ZHANG Hengxu, et al. Power system frequency response analysis based on the direct current load flow [J]. Proceedings of the CSEE, 2009, 29(34): 36-41. https://doi.org/10.13334/j.0258-8013.pcsee.2009.34.009

[11] ALIZADEH M, AMRAEE T. Adaptive scheme for local prediction of post-contingency power system frequency [J]. Electric Power Systems Research, 2014, 107(5): 240–249. https://doi.org/10.1016/j.epsr.2013.10.014

[12] BO Qibin, WANG Xiaoru, LIU Ketian. Minimum frequency prediction based on v-SVR for post-disturbance power system [J]. Electric Power Automation Equipment, 2015, 35(7): 83-88. https://doi.org/10.16081/j.issn.1006-6047.2015.07.013

[13] XU Y, DAI Y, DONG Z, et al. Extreme learning machine-based predictor for real-time frequency stability assessment of electric power systems [J]. Neural Computing and Applications, 2013, 22(1): 501-508. https://doi.org/10.1007/s00521-011-0803-3

[14] WEN Yunfeng, ZHAO Rongzhen, XIAO Youqiang, et al. Frequency safety assessment of power system based on multi-layer extreme learning machine [J]. Automation of Electric Power Systems, 2019, 43(1): 133-140. https://doi.org/10.7500/AEPS20180629012

[15] LI Guanzheng, LI Bin, WANG Shuai, et al. Dynamic frequency prediction of power system post-disturbance based on feature selection and random forest [J]. Power System Technology, 2021, 45(7): 2492-2502. https://doi.org/10.13335/j.1000-3673.pst.2021.0027

[16] ZHAO Rongzhen, WEN Yunfeng, YE Xi, et al. Research on frequency indicators evaluation of disturbance events based on improved stacked denoising autoencoders [J]. Proceedings of the CSEE, 2019, 39(14): 4081-4093. https://doi.org/10.13334/j.0258-8013.pcsee.181768

[17] HUANG Mingzeng, WEN Yunfeng, GOU Jing, et al. Maximum frequency deviation prediction method considering frequency deviation distribution and penalty cost [J]. Automation of Electric Power Systems, 2021, 45(23): 51-59. https://doi.org/10.7500/AEPS20210305006

[18] LI H, LI C, LIU Y, Maximum frequency deviation assessment with clustering based on metric learning [J]. International Journal of Electrical Power & Energy Systems, 2020, 120. https://doi.org/10.1016/j.ijepes.2020.105980

[19] JI Haoran, HOU Chunping, YANG Yang, et al. New design of self-explosive insulator detection model based on generative adversarial network [J]. Modern Electric Power, 2022, 39(5): 587-596. https://doi.org/10.19725/j.cnki.1007-2322.2021.0154

[20] CHENG Zhuo, XU Yixun, LI Zeshuang. Research on multi-time scale optimization scheduling strategy of microgrid based on improved generative adversarial network for wind and PV power scenario generation [J]. Modern Electric Power, 2023, 40. https://doi.org/10.19725/j.cnki.1007-2322.2022.0309

[21] YORINO N, ABDILLAH M, SASAKI Y. Robust power system security assessment under uncertainties using bi-level optimization [J]. IEEE Transactions on Power Systems, 2018, 33(1): 352–362. https://doi.org/10.24752/gre.1.0_302

[22] BIEIMAN L. Bagging predictors [J]. Machine Learning, 1996, 26(2): 123-140. https://doi.org/10.1023/A:1018054314350

[23] BELLET A, HABRARD A, SEBBAN M. A survey on metric learning for feature vectors and structured data. https://doi.org/10.48550/arXiv.1306.6709

[24] WEINBERGER K Q, TESAURO G. Metric learning for kernel regression [J]. Journal of Machine Learning Research, 2007, 2(5): 612-619.

[25] Yang Yanliang, Wang Xiuyun, Zeng Shuzhen, et al. Reactive power optimization based on improved gradient particle wwarm optimization algorithm [J]. Modern Electric Power, 2010, 27(4): 17-21. https://doi.org/10.19725/j.cnki.1007-2322.2010.04.004

[26] Zhao Mingyu, Xu Shiming, Gao Hui, et al. Strategy of electric vehicle emergency power supply based on fuzzy k-means algorithm [J]. Automation of Electric Power Systems, 2016, 40(05): 91-95+108. https://doi.org/10.7500/AEPS20150304001

[27] QI Sheng, SHAN Haiou, LUO Lin, et al. Application of deep feature learning with Gram's angle field for trace gas concentration identification [J]. Power System Protection and Control, 2023, 51(15): 55-65. https://doi.org/10.19783/j.cnki.pspc.221906