IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 27, 2020, accepted September 13, 2020, date of publication September 18, 2020,
date of current version September 30, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3025141

Improved Equivalent Method for Large-Scale Wind Farms Using Incremental Clustering and Key Parameters Optimization

JI HAN[1,2,3], (Student Member, IEEE), SHIHONG MIAO[1,2,3], (Member, IEEE), YAOWANG LI[1,2,3,4], (Member, IEEE), HAORAN YIN[1,2,3], DI ZHANG[1,2,3], WEICHEN YANG[1,2,3], AND QINGYU TU[1,2,3]

[1]State Key Laboratory of Advanced Electromagnetic Engineering and Technology, Huazhong University of Science and Technology, Wuhan 430074, China
[2]Hubei Electric Power Security and High Efficiency Key Laboratory, Huazhong University of Science and Technology, Wuhan 430074, China
[3]School of Electrical and Electronic Engineering, Huazhong University of Science and Technology, Wuhan 430074, China
[4]Department of Electrical Engineering, Tsinghua University, Beijing 100084, China

Corresponding author: Shihong Miao (shmiao@hust.edu.cn)

This work was supported by the National Key Research and Development Program of China under Grant 2017YFB0902600.

ABSTRACT Large-scale wind farms (WFs) generally consist of hundreds of wind turbines (WTs), and the WFs simulation model construction would be complex and even impossible if we develop each individual WT in detail. Therefore, the WFs equivalent simulation model with required accuracy is essential to be developed to explore the WFs operation characteristics. This article proposes an equivalent method for large-scale WFs using incremental clustering and key parameters optimization. Firstly, to acquire more comprehensive and distinguishable representations of WTs operation characteristics, the time series of WT active power, reactive power, voltage and current are selected as the multi-view clustering indicator (CI). Then, considering the computer memory pressures encountered by traditional clustering algorithms in dealing with large-scale WFs, a novel clustering algorithm namely multi-view incremental transfer fuzzy C-means (MVIT-FCM) is proposed, and this algorithm can process the WTs clustering problems without requiring to consider the scale of the WFs. Finally, to further increase the equivalent accuracy of the WFs equivalent simulation model, key parameters in the equivalent model are found using Sobol' criterion and then optimized using the designed Q-learning based non-dominated sorting genetic algorithm II (NSGA-II). To verify the effectiveness of the proposed method, the modified WFs system in China is utilized for case study, and the performance of the proposed model is compared with several state-of-the-art models. Simulation results show that the equivalent accuracy of the proposed model is higher when comparing with other models. Also, the proposed model has the advantage of processing the WFs equivalent problems with any scales.

INDEX TERMS Wind farm equivalence, MVIT-FCM, Q-NSGA-II, multi-view, incremental technique.

NOMENCLATURE

A. MVIT-FCM

m	weighted index, and $m \in (1,8)$
N	number of clustering samples, $j = 1, \ldots, N$
K	number of views, $k = 1, \ldots, K$
C	number of clusters, $i = 1, \ldots, C$
M	number of blocks, $l = 1, \ldots, M$
N_l	number of clustering samples in the lth block
$\boldsymbol{x}_{j,k}$	the jth sample in the kth view
$\boldsymbol{X}_k$	clustering sample matrix, and $\boldsymbol{x}_{j,k}$ is its element
$\boldsymbol{v}_{i,k}$	the ith clustering center in the kth view
$\boldsymbol{V}_k$	clustering center matrix, and $\boldsymbol{v}_{i,k}$ is its element
$\mu_{ij,k}$	membership degree of $\boldsymbol{x}_{j,k}$ to $\boldsymbol{v}_{i,k}$
$\boldsymbol{U}_k$	membership degree matrix, and $\mu_{ij,k}$ is its element
$d_{ij,k}$	distance between $\boldsymbol{x}_{j,k}$ and $\boldsymbol{v}_{i,k}$
$w_{i,k}^{(l)}$	weight coefficient in the lth block, ith clustering center and kth view
$\boldsymbol{w}_k^{(l)}$	weight coefficient matrix, and $w_{i,k}^{(l)}$ is its element

The associate editor coordinating the review of this manuscript and approving it for publication was Wen-Sheng Zhao.

$\boldsymbol{x}_{j,k}^{(l)}$	the jth sample in the lth block and kth view
$\boldsymbol{X}_k^{(l)}$	clustering sample matrix, and $\boldsymbol{x}_{j,k}^{(l)}$ is its element
$\boldsymbol{v}_{i,k}^{(l)}$	the ith clustering center in the lth block and kth view
$\boldsymbol{V}_k^{(l)}$	clustering center matrix, and $\boldsymbol{v}_{i,k}^{(l)}$ is its element
$\mu_{ij,k}^{(l)}$	membership degree of $\boldsymbol{x}_{j,k}$ to $\boldsymbol{v}_{i,k}$ in the lth block
$\boldsymbol{U}_k^{(l)}$	membership degree matrix, and $\mu_{ij,k}^{(l)}$ is its element
$\alpha_{j,k}$	importance coefficient of $x_{j,k}$ in rewarding membership degree
$\beta_{j,k}$	penalty coefficient used to weaken $\alpha_{j,k}$
λ	regularization parameter, and $\lambda>0$
γ	trade-off factor, and $\gamma \in (0,1)$
$\hat{\mu}_{ij}$	overall membership

B. TRANSFERRED DATA SELECTION METHOD

W_i	wind speed of the ith WT
$\overline{W}_j$	wind speed of the jth LVRT experiment
$V_{\mathrm{dip},i}$	voltage dip value of the ith WT
$\overline{V}_{\mathrm{dip},j}$	voltage dip value of the jth LVRT experiment
w_1,w_2	weight coefficient, and $w_1 + w_2 = 1$
s_{ij}	comprehensive similarity coefficient between the ith WT in WF and the jth LVRT experiment
n_{sim}	a positive integer
Ex_i	the n_{sim} LVRT experiments that are most similar to the ith WT
$\mathrm{Ex}_{\mathrm{tr}}$	overall LVRT experiments transferred to target domain

C. KEY PARAMETERS OPTIMIZATION OF WF EQUIVALENT MODEL

S_i	first-order sensitivity coefficient
$S_{\mathrm{T}i}$	total sensitivity coefficient
ε_1 *and* ε_{T}	threshold value
f_P, f_Q, f_V, f_I	sub-objective functions, representing equivalent precision of active power, reactive power, voltage and current, respectively
ΔP_t^n, ΔQ_t^n, ΔV_t^n, ΔI_t^n	difference between equivalent WF and actual WF in active power, reactive power, voltage and current, respectively
y_1-y_{56}	parameters to be optimized in the equivalent unit
$\boldsymbol{y}_{\mathrm{key}}$	vector of key parameters to be optimized
$\boldsymbol{y}_{\mathrm{key,min}}$	optimized lower limits of the key parameters
$\boldsymbol{y}_{\mathrm{key,max}}$	optimized upper limits of the key parameters
Ω_{typ}	typical LVRT circumstances
N_{typ}	number of LVRT circumstances in Ω_{typ}
T	points number of the time series
$\boldsymbol{Q}$, $\boldsymbol{\Delta Q}$	knowledge matrix and knowledge increment
$\boldsymbol{A}$	action space
s and a	state and action
$R(s_{im}^{kj}, s^{k+1,j}, \alpha_{im}^{kj})$	reward function of a transition from state s^k to s^{k+1} under a selected action a^k
QI	number of key parameters, i=1,2,…,QI
QM_i	binary bit string length of key parameter i, m=1,2,…,QM_i
QJ	number of individuals, j=1,2,…,QJ
QK	maximum iteration times, k=1,2,…,QK
α	knowledge learning factor
γ	discount factor
r_0	a random value with the range of [0,1]
ε	exploitation rate
a_{rand}	a random action
W	a positive constant
$\boldsymbol{SA}_i^{\mathbf{BEST}}$	state-action pairs set of the best individual at the k-th iteration of i, $i \in P, Q, V, I$
$f_i^{\min}, f_i^{\max}$	maximum and minimum values of the sub-objective function i, $i \in P, Q, V, I$

D. ERROR METRIC

e_{RMSE}	root mean square error
e_{MAE}	mean absolute error
e_{MAPE}	mean absolute percentage error
X	actual value of the time series
$\widehat{X}$	predicted value of the time series
L	length of time series X and $\widehat{X}$

I. INTRODUCTION

Wind power is one of the most mature new energy power generation techniques. Until 2020, the installed capacity of wind turbines (WTs) would be 2.1 billion kW in China [1]. In power systems, the short-circuit faults occur frequently, thereby resulting in voltage dip circumstances in the output port of wind farms (WFs) systems. To ensure the safe operation of power systems, WFs are required to have low voltage ride through (LVRT) ability and are capable of providing the reactive power to help maintain system stability when voltage drops. To study LVRT dynamic process, the dynamic WFs model is necessary to establish [2]. Yet, there might be hundreds of WTs in the large-scale WFs, and it is not feasible to model each WT in detail because of the huge model size and simulation time [3]. Thus, it is critical to develop the WFs equivalent model for LVRT analysis on the basis of reasonable reduction [4].

In general, there are two types of WFs equivalent methods: single-machine equivalence (SME) and multi-machines equivalence (MME), where MME-based methods possess higher accuracy and border applications [5]. Three steps are generally taken in MME.

Firstly, the clustering indicator (CI) needs to be selected to evaluate the differences of the operation characteristics among different WTs. Until now, the researchers have explored to use wind speed [3], pitch angle [4], rotor speed [5], active power [6] as the CIs. Secondly, based on the selected CI, WTs clustering would be implemented. This

step aims to cluster the WTs with the same or similar operation characteristics into the same group, and pushes the WTs operating quite different to different groups. Generally, clustering algorithms are commonly used to achieve this division, and fuzzy C-means (FCM) clustering [6], K-means clustering [7], hierarchical clustering [8] have already been fully investigated. Based on the above two steps, the WTs have been divided into several groups. Each group is called as one equivalent unit, which consists of equivalent WT, transformer and collector network. Thus, thirdly, the equivalent parameters of the equivalent units need to be calculated, including body parameters and control parameters. The capacity weighted method is most commonly used [4], and there are also researchers using equal power loss method [9] for collector network parameters equivalence.

Although favorable equivalent results are achieved using the above methods, some limitations are still existed. Firstly, the CIs in most of the researches focus on single-view. For example, the active power time series is used as CI in [10]. Yet, the active power in some WTs might be similar, but their reactive power might perform quite different if they are set different reactive power reference values. Apparently, using active power as the CI, the final equivalent accuracy of active power would be high, but the accuracy of reactive power cannot be ensured. Thus, in [11], [12], we have firstly proposed to use multi-view time series-based CI, so as to increase the equivalent accuracy in comprehensive aspects. However, our methods face challenges when processing large-scale WFs consisting of hundreds or more WTs. This is because the dimensionality of the time series-based CIs is quite high. When the WTs number is large, the size of the clustering-used data would be very huge, resulting in severe computer memory pressure and even computer freezing phenomenon when processing these large-size data samples.

Another limitation lies in the equivalent parameters calculation step. Although capacity weighted method has the advantage of simple calculation and can acquire well-performed results, it only focuses on the calculation of body parameters, and generally omits the calculation of control parameters and sets them to default values [13]. Yet, with LVRT, there are many dynamic response processes, which are mainly influenced by control parameters. Thus, the default values setting for control parameters would limit the further improvement of equivalent accuracy to some extent.

To address the above two limitations in large-scale WFs processing and the equivalent parameters calculation, this article proposes an equivalent method for large-scale WFs using incremental clustering and key parameters optimization, main contributions of this article can be summarized as follows.

(1) In order to conveniently process the clustering problems of large-scale WFs, incremental technique [14] is introduced to the calculation process of multi-view transfer FCM (MVT-FCM) [11] algorithm. We name the new clustering algorithm as multi-view incremental transfer FCM (MVIT-FCM). Incremental technique divides clustering-used data into multiple small data blocks, and different blocks are coupled together in the clustering process. With the aid of incremental technique, the calculation would not be limited by the data size, increasing the feasibility of proposed equivalent method when processing large-scale WFs.

(2) In the equivalent parameters calculation step, both control and body parameters of the equivalent units are optimized using Q-learning based non-dominated sorting genetic algorithm II (NSGA-II). Considering the number of the to-be-optimized parameters is very large, which is difficult to solve for optimization problems, Sobol'-based criterion [15] is designed to find the key parameters to be optimized. Then, these key parameters are optimized using the designed Q-NSGA-II, which is an improved algorithm of NSGA-II [16] by introducing Q-learning [17] in it. This equivalent parameters calculation method improves the accuracy of the equivalent WF further, and has high calculation efficiency.

The paper is organized as follows. The multi-view CI consideration and the discussion are presented in Section II. In Section III, MVIT-FCM is proposed. In Section IV, the Q-NSGA-II based equivalent parameter calculation method after WTs clustering is given. Section V provides the overall WFs equivalent process. The case study is provided in Section VI. Section VII concludes the paper.

II. MULTI-VIEW CI SELECTION

In the previous equivalent studies of WFs, the time series-based CI often focuses on single-view, such as active power time series. Yet, the reactive power (or other physical quantities) of two WTs would be different if they are set to different reactive power reference values or their product types are different, even though they have the same active power. In this article, a WT model is built in Matlab / Simulink, as shown in Fig. 1 (a). Set different reactive power reference values for the two WTs, and apply them the same wind speed disturbance. Fig. 1 (b)-(c) shows the active and reactive power curves of the two WTs. From this figure, we can see the active power curves of the two WTs are basically the same, but the reactive power curves are quite different. Therefore, using single-view active power time series as the evaluation of the external characteristics of the WTs is somewhat limited, and cannot comprehensively reflect the output characteristics of the WTs in multi-view perspectives [11], [12]. Therefore, the selected CI should have the ability in describing the characteristics of the WTs completely. This article selects the time series of the active power, reactive power, voltage and current at the outlet of the WTs as the CI.

III. WTs CLUSTERING BASED ON MVIT-FCM

The fundamental algorithms of MVIT-FCM, including multi-view FCM (MV-FCM) and MVT-FCM, are reviewed firstly. Then, MVIT-FCM algorithm is proposed and its

TABLE 1. Review of MV-FCM and MVT-FCM.

Algorithm	MV-FCM	MVT-FCM
Optimization objective	$\min\left\{\sum_{k=1}^{K}\sum_{j=1}^{N}\sum_{i=1}^{C}[\mu_{ij,k}^{m}d_{ij,k}^{2}+\alpha_{j,k}\mu_{ij,k}(1-\mu_{ij,k}^{m-1}) -\beta_{j,k}\mu_{ij,k}(1-\mu_{ij,k}^{m-1})]\right\}$	$\min\left\{\sum_{k=1}^{K}\sum_{j=1}^{N}\sum_{i=1}^{C}[\gamma\mu_{ij,k}^{m}d_{ij,k}^{2}+(1-\gamma)\bar{\mu}_{ij,k}^{m}d_{ij,k}^{2}+\alpha_{j,k}\mu_{ij,k}(1-\mu_{ij,k}^{m-1}) -\beta_{j,k}\mu_{ij,k}(1-\mu_{ij,k}^{m-1})]+\lambda\sum_{k=1}^{K}\sum_{i=1}^{C}\left\|\boldsymbol{v}_{i,k}-\bar{\boldsymbol{v}}_{i,k}\right\|_2^2\right\}$
Constraints	$\sum_{i=1}^{C}\mu_{ij,k}=1;\mu_{ij,k}\in[0,1];1\le i\le C;1\le j\le N;1\le k\le K$	$\sum_{i=1}^{C}\mu_{ij,k}=1;\mu_{ij,k}\in[0,1];1\le i\le C;1\le j\le N;1\le k\le K$
Updating equations	$\mu_{ij,k}=1\Big/\sum_{i'=1}^{C}\left[\frac{\gamma d_{ij,k}^{2}-(\alpha_{j,k}-\beta_{j,k})}{\gamma d_{ij,k}^{2}-(\alpha_{j,k}-\beta_{j,k})}\right]^{\frac{1}{m-1}}$ $\boldsymbol{v}_{i,k}=\frac{\sum_{j=1}^{N}\mu_{ij,k}^{m}\boldsymbol{x}_{j,k}}{\sum_{j=1}^{N}\mu_{ij,k}^{m}}$	$\mu_{ij,k}=1\Big/\sum_{i'=1}^{C}\left[\frac{\gamma d_{ij,k}^{2}-(\alpha_{j,k}-\beta_{j,k})}{\gamma d_{ij,k}^{2}-(\alpha_{j,k}-\beta_{j,k})}\right]^{\frac{1}{m-1}}$ $\boldsymbol{v}_{i,k}=\frac{\lambda\bar{\boldsymbol{v}}_{i,k}+\sum_{j=1}^{N}[\gamma\mu_{ij,k}^{m}+(1-\gamma)\bar{\mu}_{ij,k}^{m}]\boldsymbol{x}_{j,k}}{\lambda+\sum_{j=1}^{N}[\gamma\mu_{ij,k}^{m}+(1-\gamma)\bar{\mu}_{ij,k}^{m}]}$
Pseudocode	Input: $C,m,\boldsymbol{w}_k,\boldsymbol{X}_k$ Output: $\boldsymbol{U}_k(\boldsymbol{\mu}_{ij,k}),\boldsymbol{V}_k(\boldsymbol{v}_{i,k})$ while $\max_{1\le i\le C,1\le k\le K}\left\|\boldsymbol{v}_{i,k,\text{new}}-\boldsymbol{v}_{i,k,\text{old}}\right\|_2^2>\boldsymbol{\varepsilon}$ $\boldsymbol{U}_k$ updation $\boldsymbol{V}_k$ updation end	Input: $C,m,\boldsymbol{w}_k,\boldsymbol{X}_k,\bar{\boldsymbol{U}}_k,\bar{\boldsymbol{V}}_k$ Output: $\boldsymbol{U}_k(\boldsymbol{\mu}_{ij,k}),\boldsymbol{V}_k(\boldsymbol{v}_{i,k})$ while $\max_{1\le i\le C,1\le k\le K}\left\|\boldsymbol{v}_{i,k,\text{new}}-\boldsymbol{v}_{i,k,\text{old}}\right\|_2^2>\boldsymbol{\varepsilon}$ $\boldsymbol{U}_k$ updation $\boldsymbol{V}_k$ updation end

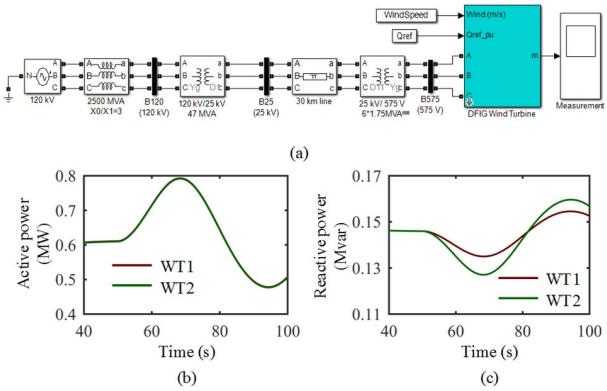


FIGURE 1. Model and output characteristics of the WT. (a) simulation model, (b) active power and (b) reactive power.

calculation process is provided. Finally, the source domain data selection method to help the WTs clustering is given.

A. FUNDAMENTAL ALGORITHM: MV-FCM AND MVT-FCM

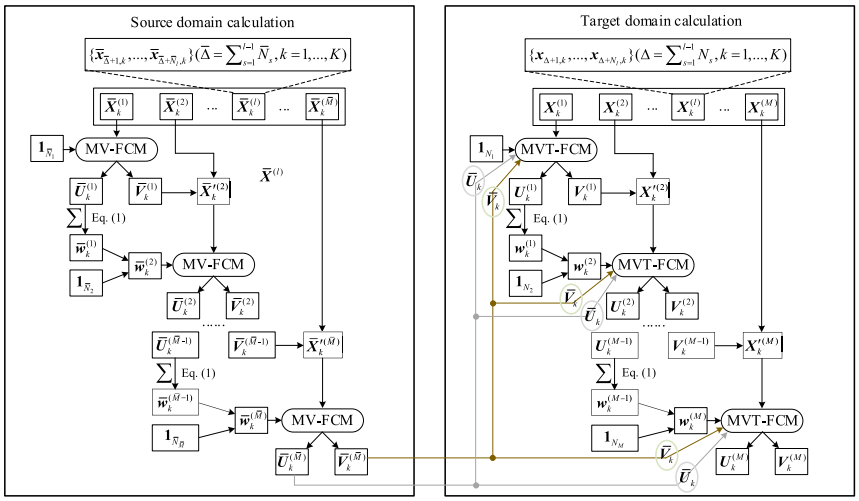
MV-FCM [18] belongs to multi-view clustering algorithm, which can reward or suppress the corresponding fuzzy membership degrees for the current view, and combines the clustering results in each view into global clustering result. In [11], we proposed a new clustering algorithm for WTs clustering, namely MVT-FCM, which is an improved algorithm based on MV-FCM. In MVT-FCM, there are two domains of data samples, i.e., source domain data and target domain data. The MVT-FCM algorithm transfers the beneficial knowledge of source domain into the target domain to guide the clustering process, and this auxiliary guidance significantly enhances the stability and accuracy of WTs clustering.

Before the presentation of MVIT-FCM, we firstly review the MV-FCM and the proposed MVT-FCM. We declare, if two variables use the same symbol, the variable with top horizontal line is from source domain, and that without top horizontal line is from target domain, the meanings of two variables are totally the same. Due to the page limitation, we just list the optimization objective, constraints, updating equations and pseudocode, which can be found in TABLE 1.

With MVT-FCM algorithm, $\sum_{k=1}^{K}\sum_{i=1}^{C}\left\|\boldsymbol{v}_{i,k}-\bar{\boldsymbol{v}}_{i,k}\right\|_2^2$ is the first transfer rule which aims to transfer the clustering center of source domain to the target domain, and $\sum_{k=1}^{K}\sum_{j=1}^{N}\sum_{i=1}^{C}\bar{\mu}_{ij,k}^{m}d_{ij,k}^{2}$ is the second transfer rule which aims to transfer the membership of source domain to the target domain. Based on these two transfer rules, the algorithm can effectively deal with the problem of low clustering accuracy and instability because of the rare data amount in target domain [19].

B. ADVANCED ALGORITHM: MVIT-FCM

Further, incremental technique is applied to MVT-FCM. We name the new algorithm as MVIT-FCM. Incremental technique firstly divides the samples into multiple small sets, each of which is called as a block. Then, MVT-FCM is applied to each block, and different blocks are coupled with weight coefficient and updating clustering center. Because of the blocks division, the computer can deal with the separated blocks sequentially. Generally, the blocks division is arbitrary, and the data-size of one block can be much smaller than the whole dataset. Thus, MVIT-FCM has the ability of processing the large-size dataset.



(a)

Input: $C, m, \boldsymbol{w}_k, \boldsymbol{X}_k, N_l, \overline{\boldsymbol{w}}_k, \overline{\boldsymbol{X}}_k, \overline{N}_l$
Output: $\boldsymbol{U}_k, \boldsymbol{V}_k$

Divide $\overline{\boldsymbol{X}}_k$ into $\overline{M}$ blocks,
which is $\{\overline{\boldsymbol{X}}_k^{(1)}, ..., \overline{\boldsymbol{X}}_k^{(\overline{M})}\}$
$\overline{\boldsymbol{w}}_k^{(1)} = \mathbf{1}_{\overline{N}_1}$
$\overline{\boldsymbol{U}}_k^{(1)}, \overline{\boldsymbol{V}}_k^{(1)} = \text{MV-FCM}(C, m, \overline{\boldsymbol{w}}_k^{(1)}, \overline{\boldsymbol{X}}_k^{(1)})$
for $l = 2$ to $\overline{M}$
  Calculate $\overline{\boldsymbol{w}}_k^{(l-1)}$ using (1)
  $\overline{\boldsymbol{w}}_k^{(l)} = \{\overline{\boldsymbol{w}}_k^{(l-1)} \bigcup \mathbf{1}_{\overline{N}_l}\}$
  $\overline{\boldsymbol{X}}_k'^{(l)} = \{\overline{\boldsymbol{X}}_k^{(l)} \bigcup \overline{\boldsymbol{V}}_k^{(l-1)}\}$
  $\overline{\boldsymbol{U}}_k^{(l)}, \overline{\boldsymbol{V}}_k^{(l)} = \text{MV-FCM}(C, m, \overline{\boldsymbol{w}}_k^{(l)}, \overline{\boldsymbol{X}}_k'^{(l)})$
end
$\overline{\boldsymbol{U}}_k = \overline{\boldsymbol{U}}_k^{(\overline{M})}, \overline{\boldsymbol{V}}_k = \overline{\boldsymbol{V}}_k^{(\overline{M})}$

Divide $\boldsymbol{X}_k$ into M blocks,
which is $\{\boldsymbol{X}_k^{(1)}, ..., \boldsymbol{X}_k^{(M)}\}$
$\boldsymbol{w}_k^{(1)} = \mathbf{1}_{N_1}$
$\boldsymbol{U}_k^{(1)}, \boldsymbol{V}_k^{(1)} = \text{MVT-FCM}(C, m, \boldsymbol{w}_k^{(1)}, \boldsymbol{X}_k^{(1)}, \overline{\boldsymbol{U}}_k, \overline{\boldsymbol{V}}_k)$
for $l = 2$ to M
  Calculate $\boldsymbol{w}_k^{(l-1)}$ using (1)
  $\boldsymbol{w}_k^{(l)} = \{\boldsymbol{w}_k^{(l-1)} \bigcup \mathbf{1}_{N_l}\}$
  $\boldsymbol{X}_k'^{(l)} = \{\boldsymbol{X}_k^{(l)} \bigcup \boldsymbol{V}_k^{(l-1)}\}$
  $\boldsymbol{U}_k^{(l)}, \boldsymbol{V}_k^{(l)} = \text{MVT-FCM}(C, m, \boldsymbol{w}_k^{(l)}, \boldsymbol{X}_k'^{(l)}, \overline{\boldsymbol{U}}_k, \overline{\boldsymbol{V}}_k)$
end
$\boldsymbol{U}_k = \boldsymbol{U}_k^{(M)}, \boldsymbol{V}_k = \boldsymbol{V}_k^{(M)}$

(b)

FIGURE 2. Clustering process of MVIT-FCM. (a) the flow chart, and (b) the pseudocode.

The clustering process of MVIT-FCM is shown in Fig. 2. It should be noted that the expression of Eq. (1) is as follows.

$$w_{i,k}^{(l)} = \sum_{j=1}^{N_l} \mu_{ij,k}^{(l)} \quad (i = 1, 2, \ldots, C; 1 \le k \le K) \tag{1}$$

In Fig. 2, the membership and clustering center matrices of source and target domains are calculated, respectively. With the calculation of $\overline{\boldsymbol{U}}_k$ and $\overline{\boldsymbol{V}}_k$ in source domain, $\overline{\boldsymbol{X}}_k$ is divided into $\overline{M}$ blocks, namely $\overline{\boldsymbol{X}}_k^{(l)}$, where $l = 2, \ldots, \overline{M}$. Then, $\overline{\boldsymbol{U}}_k^{(1)}$ and $\overline{\boldsymbol{V}}_k^{(1)}$ can be acquired using the 2nd-3rd rows of the pseudocode in Fig. 2. Based on this, the *for* loop begins to calculate $\overline{\boldsymbol{U}}_k^{(l)}$ and $\overline{\boldsymbol{V}}_k^{(l)}$ in turns using the 5th-8th rows of the pseudocode, where $l = 2, \ldots, \overline{M}$. Finally, set $\overline{\boldsymbol{U}}_k = \overline{\boldsymbol{U}}_k^{(\overline{M})}$ and $\overline{\boldsymbol{V}}_k = \overline{\boldsymbol{V}}_k^{(\overline{M})}$.

The $\overline{\boldsymbol{U}}_k$ and $\overline{\boldsymbol{V}}_k$ terms are further transferred to target domain to help the calculation of $\boldsymbol{U}_k$ and $\boldsymbol{V}_k$, whose calculation process is similar to that of $\overline{\boldsymbol{U}}_k$ and $\overline{\boldsymbol{V}}_k$. Specifically, $\boldsymbol{X}_k$ is divided into M blocks, namely $\boldsymbol{X}_k^{(l)}$, where $l = 2, \ldots, M$. Then, $\boldsymbol{U}_k^{(1)}$ and $\boldsymbol{V}_k^{(1)}$ can be acquired using the 12th-13th rows of the pseudocode in Fig. 2. Based on this, the *for* loop begins to calculate $\boldsymbol{U}_k^{(l)}$ and $\boldsymbol{V}_k^{(l)}$ in turns using the 15th-18th rows of the pseudocode, where $l = 2, \ldots, M$. Finally, set $\boldsymbol{U}_k = \boldsymbol{U}_k^{(M)}$ and $\boldsymbol{V}_k = \boldsymbol{V}_k^{(M)}$.

After obtaining $\boldsymbol{U}_k$, the overall membership can be calculated using (2)

$$\hat{\mu}_{ij} = \sqrt[K]{\prod_{k=1}^{K} \mu_{ij,k}} \tag{2}$$

C. TRANSFERRED DATA SELECTION METHOD

In this article, the time series of active power, reactive power, voltage and current in each WT reflect the operation characteristics of the WTs. We select these time series as target domain data. With the transfer learning-based MVIT-FCM algorithm, the source domain data selection is essentially important. If the data distribution of the two domains are not

similar enough, the introducing of the source domain data would worsen the target domain calculation result [20].

To acquire the source domain data suitable and helpful for clustering process, abundant LVRT simulation experiments are performed under different wind speed and voltage dip values, and we would select the source domain data from these LVRT experiments according the similarity between two domains. Fig. 3 and Fig. 4 show the active power and reactive power curves of the WT at various wind speed and voltage dip values.

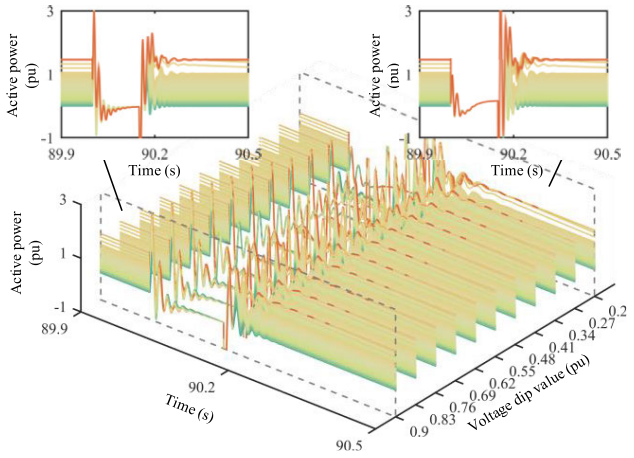


FIGURE 3. Active power curves of WT at various LVRT conditions.

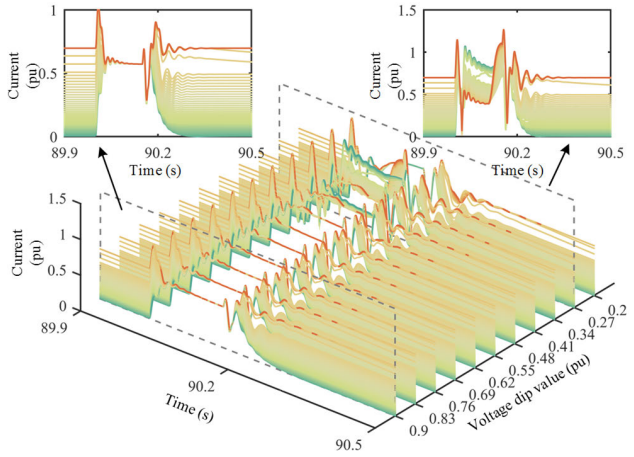


FIGURE 4. Current curves of WT at various LVRT conditions.

For simplicity, we consider the wind speed and voltage dip values to judge the similarity of the data in source and target domain, and only transfer the source domain samples that are most similar to target domain. Define the comprehensive similarity coefficient, which is as follows

$$s_{ij} = w_1 \left| \frac{W_i - \overline{W}_j}{\overline{W}_j} \right| + w_2 \left| \frac{V_{\mathrm{dip},i} - \overline{V}_{\mathrm{dip},j}}{\overline{V}_{\mathrm{dip},j}} \right| \tag{3}$$

Apparently, the smaller value of s_{ij}, the more similar of the operation characteristics between the ith WT and the jth LVRT experiment. As for the ith WT, select the n_{sim} LVRT experiments that are most similar to it, namely Ex_i. Finally, the time series of active power, reactive power, voltage and current corresponding to the following LVRT experiments are transferred to target domain.

$$\mathrm{Ex}_{\mathrm{tr}} = \bigcup_{i=1}^{N} \mathrm{Ex}_i \tag{4}$$

IV. KEY PARAMETERS OPTIMIZATION OF WF EQUIVALENT MODEL

After WTs clustering using the MVIT-FCM algorithm, the topology of the equivalent WF is shown in Fig. 5. The body and control parameters of the equivalent unit, including equivalent WT, equivalent transformer and equivalent collector network, need to be calculated. The Sobol'-based method for selecting key parameters to be optimized is given firstly. Then, the key parameters optimization model is provided. Finally, we provide the optimization model solving method, i.e., Q-NSGA-II, which is an improved algorithm of NSGA-II by introducing Q-learning in it to increase the algorithm solving efficiency.

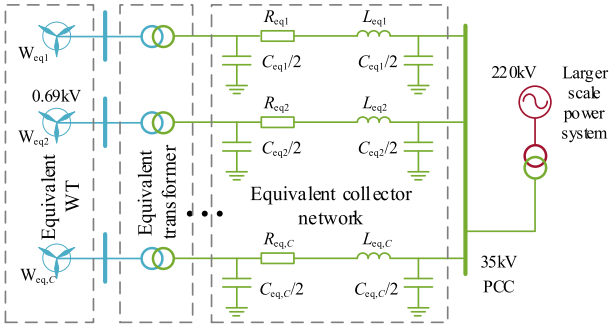


FIGURE 5. The layout of the equivalent wind farms.

A. SOBOL'-BASED METHOD FOR SELECTING KEY PARAMETERS

According to statistics, the number of body and control parameters in one equivalent unit (including equivalent WT, equivalent transformer and equivalent collector network) is 56. The meaning of the parameters are listed in Appendix A, and we denote these parameters as $y_1 - y_{56}$. In the WFs equivalent model, there would be C equivalent WTs, so there are $C \times 56$ parameters need to be optimized, which is quite hard to solve for the optimization problems. Therefore, the key parameters, which influence the LVRT process more, need to be selected.

Sobol' criterion based for selecting key parameters is provided. Sobol' is a tool for analyzing the effect or sensitivity of input parameters on system output. Two sensitivity coefficients are commonly-used: first-order sensitivity coefficient (S_i), which is used to estimate the influence of input parameter i to output; total sensitivity coefficient ($S_{\mathrm{T}i}$), which is used to estimate the influence of input parameter i and other parameters together to output. Due to the page limitation,

the calculation method of S_i and S_{Ti} can be found in [15]. According to the feature of Sobol', the method for selecting key parameters are as follows.

(1) As for the 56 parameters in each equivalent WT, set the intervals, 20% up and down of the original values, as the sampling intervals.
(2) Random sampling method is used to sample the points from the intervals. Using the sampled values, LVRT experiments are developed, and the time series of WT active power, reactive power, voltage and current are acquired.
(3) Based on the time series results of the LVRT experiments, S_i and S_{Ti} are calculated for each parameter. Set the threshold values ε_1 and ε_T, and choose the parameters with $S_i > \varepsilon_1$ and $S_{Ti} > \varepsilon_T$ as the key parameters.

B. KEY PARAMETERS OPTIMIZATION MODEL

Based on the selected key parameters, the key parameters optimization model is constructed, and they are optimized using Q-NSGA-II. The objective function considers the equivalent precision of active power, reactive power, voltage and current at PCC. Further, to ensure the optimization results can possess high accuracy in various LVRT circumstances, several typical LVRT circumstances are considered in the optimization process. The objective function of the key parameters optimization model can be expressed as

$$\begin{aligned}
\min F(y) &= \min[f_P(\boldsymbol{y}_{\text{key}}), f_Q(\boldsymbol{y}_{\text{key}}), f_V(\boldsymbol{y}_{\text{key}}), f_I(\boldsymbol{y}_{\text{key}})] \\
f_P(\boldsymbol{y}_{\text{key}}) &= \frac{1}{T \times N_{\text{typ}}} \sqrt{\sum\nolimits_{n\in\Omega_{\text{typ}}} \sum\nolimits_{t=1}^{T} (\Delta P_t^n)^2} \\
f_Q(\boldsymbol{y}_{\text{key}}) &= \frac{1}{T \times N_{\text{typ}}} \sqrt{\sum\nolimits_{n\in\Omega_{\text{typ}}} \sum\nolimits_{t=1}^{T} (\Delta Q_t^n)^2} \\
f_V(\boldsymbol{y}_{\text{key}}) &= \frac{1}{T \times N_{\text{typ}}} \sqrt{\sum\nolimits_{n\in\Omega_{\text{typ}}} \sum\nolimits_{t=1}^{T} (\Delta V_t^n)^2} \\
f_I(\boldsymbol{y}_{\text{key}}) &= \frac{1}{T \times N_{\text{typ}}} \sqrt{\sum\nolimits_{n\in\Omega_{\text{typ}}} \sum\nolimits_{t=1}^{T} (\Delta I_t^n)^2}
\end{aligned} \tag{5}$$

where the Ω_{typ} include the typical LVRT circumstances shown in TABLE 2. According to TABLE 2, there are 9 LVRT circumstances. Thus, $N_{\text{typ}} = 9$.

TABLE 2. Typical LVRT circumstances.

Number	Wind speed	Voltage dip
1	10m/s	0.2pu
2	15m/s	0.2pu
3	20m/s	0.2pu
4	10m/s	0.4pu
5	15m/s	0.4pu
6	20m/s	0.4pu
7	10m/s	0.6pu
8	15m/s	0.6pu
9	20m/s	0.6pu

The constraint condition of the key parameters optimization model is

$$\boldsymbol{y}_{\text{key,min}} \le \boldsymbol{y}_{\text{key}} \le \boldsymbol{y}_{\text{key,max}} \tag{6}$$

C. THE SOLVING METHOD: Q-NSGA-II

Eq. (5) has several objective functions, and can be solved using multi-objective optimization algorithms, such as NSGA-II algorithm. With NSGA-II, the way that individuals generation in NSGA-II include crossover and mutation, where mutation is a purely random behavior with no selectivity. To increase the optimization efficiency, Q-learning is introduced to NSGA-II for increasing algorithm performance. We denote this method as Q-NSGA-II. The basic principle of Q-NSGA-II is to replace the random mutation in NSGA-II to the Q-learning based intelligent mutation. The detail of NSGA-II can be found in [16], this article only introduces the Q-learning based intelligent mutation.

The Q-learning based intelligent mutation would not be random behaviors. This article design the mutation behavior using the following ε-greedy rule based action selection, which yields

$$a_{im}^{kj} = \begin{cases} \arg\max\limits_{a_{im}\in \boldsymbol{A}_{im}} \boldsymbol{Q}_{im}^{k}(s_{im}^{kj}, a_{im}) & r_0 \le \varepsilon \\ a_{\text{rand}} & r_0 > \varepsilon \end{cases} \tag{7}$$

From Eq. (7), the mutation behavior would be mainly influenced by the knowledge matrix $\boldsymbol{Q}$, whose updating equation is

$$\begin{cases} \boldsymbol{Q}_{im}^{k+1}(s_{im}^{kj}, a_{im}^{kj}) = \boldsymbol{Q}_{im}^{k}(s_{im}^{kj}, a_{im}^{kj}) + \alpha \Delta \boldsymbol{Q}_{im}^{k} \\ \Delta \boldsymbol{Q}_{im}^{k} = R_{im}^{j}(s_{im}^{kj}, s_{im}^{k+1,j}, a_{im}^{kj}) \\ \qquad\qquad + \gamma \max\limits_{a_{im}\in \boldsymbol{A}_{im}} \boldsymbol{Q}_{im}^{k}(s_{im}^{k+1,j}, a_{im}) - \boldsymbol{Q}_{im}^{k}(s_{im}^{kj}, a_{im}^{kj}) \end{cases} \tag{8}$$

The updating process of Eq. (8) is shown in Fig. 6. In Fig. 6, y_1, y_2, $y_6, \ldots, y_{49}$ are the selected key parameters, i.e., $\boldsymbol{y}_{\text{key}}$, and we would explain their selection process and result in Section VI-B. Specially, it should be

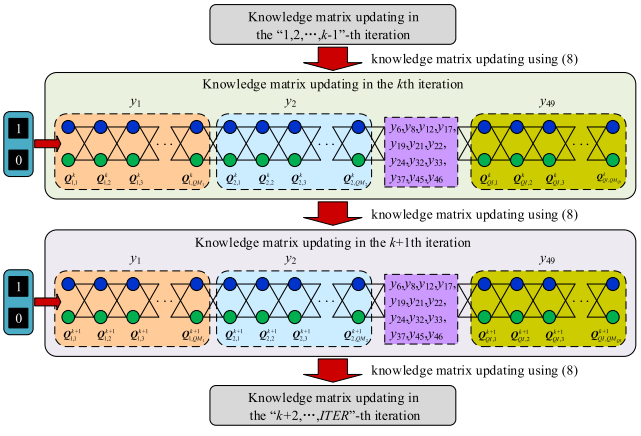


FIGURE 6. Binary state-action chain based knowledge matrix.

noted that the key parameters belong to continuous variables, but Q-learning generally processes discrete variables. Thus, the binary state-action chain [17] is adopted for the discretization of the continuous variables, where each key parameter corresponds to several knowledge matrices (For example, there are QM_1 knowledge matrices for y_1).

In addition, the reward function R in Eq. (8) is designed as follows

$$R(s_{im}^{kj}, s_{im}^{k+1,j}, a_{im}^{kj}) = \begin{cases} \dfrac{W}{\sum\limits_{i\in P,Q,V,I}\left(\min\limits_{j=1,2,\ldots,QJ} f_i^{kj}\right)} & (s_{im}^{kj}, a_{im}^{kj}) \in \boldsymbol{SA}_i^{\mathrm{BEST}} \\ 0 & (s_{im}^{kj}, a_{im}^{kj}) \notin \boldsymbol{SA}_i^{\mathrm{BEST}} \end{cases} \tag{9}$$

After the optimization of Q-NSGA-II, we would acquire some Pareto optimal solutions, and we should select one as the final optimization result of the key parameters. In this article, the fuzzy set theory [21] is adopted to determine the optimal compromise solution, and the Pareto optimal solution corresponding to the maximum h value is determined as the final optimization result of the key parameters.

$$h = \sum_{i\in P,Q,V,I}\left(\frac{f_i^{\max} - f_i}{f_i^{\max} - f_i^{\min}}\right) \tag{10}$$

V. WFs EQUIVALENT PROCESS

In this article, the time series of WT active power, reactive power, voltage and current are selected as the multi-view CI, which are input to MVIT-FCM to achieve the clustering of WTs. Based on this, the key parameters are optimized using Q-NSGA-II. The WFs equivalent process are listed as follows.

(1) Given the wind speed and PCC voltage dip value, the wind speed on each WT can be calculate using wake effect model [22], and the voltage dip value on each WT can be calculate using the power flow calculation method in Appendix B.
(2) According to the wind speed and voltage dip values of each WT, the active power, reactive power, voltage and current time series of each WT can be acquired through simulation. These time series are used as the target domain data to be clustered.
(3) Select the appropriate n_{sim}, and use Eq. (3) and Eq. (4) to acquire the transferred data, which are used as the source domain data to be clustered.
(4) For MVIT-FCM, input the source domain and the target domain data. Using Fig. 2, the clustering result based on the proposed MVIT-FCM algorithm can be acquired.
(5) According to the calculation process of Sobol', choose the key parameters.
(6) Construct the key parameters optimization model using Eq. (5). Then, optimize the key parameters using Q-NSGA-II. With non-key parameters, we also use the capacity weighted method to calculate them.

In the above WFs equivalent process, procedure (1), (2) and (3) collect the clustered data samples. Procedure (4) achieves WTs clustering using MVIT-FCM. Procedure (5) and (6) calculate the equivalent parameters. It should be noted that the static power flow calculation in procedure (1) is just used to determine the voltage dip values (initial conditions) on the WTs. Using these initial conditions, the dynamic LVRT equivalent experiments can be further performed. With the proposed MVIT-FCM, it can be encapsulated into a software package and the users just need to acquire the algorithm inputs using procedure (1), (2) and (3), which are regular steps in WFs equivalence. With the equivalent parameters calculation, it only needs to be carried out once and the calculation results can be reused because we have considered typical LVRT circumstances in the optimization process. Thus, the proposed method does not require much manual experiences and the usage is quite simple, proving it a powerful tool in engineering applications.

VI. SIMULATION RESULTS

The test WFs system and WTs clustering results are given firstly. Then, the Sobol' criterion based key parameters selection process and their optimization results are provided. Further, the accuracy of the proposed WFs equivalent method is tested through comparing with several other methods. Next, we test the performance of MVIT-FCM when processing different scale WFs. Finally, we test the optimization efficiency of Q-NSGA-II and verify the superiority of the equivalent parameters calculation method.

A. THE TEST WFs SYSTEM AND WTs CLUSTERING RESULTS

In this article, the simulation is carried out based on an actual system in East Inner Mongolia of China and a certain degree of expansion. The layout of the test case is shown in Fig. 7. In this system, the wind power hub station consists of 5WFs, totally 200 WTs, and all of the WTs connect to collector lines by 0.69/35kV box transformers and access to external system by 35/220kV transformer. Specifically, the detailed layout of the WTs in the 1st WF is shown in Fig. 7. With the 2nd-5th WFs, their WTs layout are modified on the basis of the 1st WF, and this article would not show them in detail.

Set three-phase short-circuit fault at PCC with voltage dip value 0.2pu, and the voltage dip value on WTs can be calculated through power flow calculation method shown in Appendix B. When fault occurs, the wind speed is 20m/s, and the wind speed of each WT is calculate using wake effect. Using these wind speed and voltage dip values, the time series of active power, reactive power, voltage and current in each WT are simulated, and these time series are selected as the of target domain samples. Further, select $n_{\mathrm{sim}} = 10$, and the LVRT experiments that are most similar to WTs in WF are obtained using equations (3) and (4), totally $10 \times 200 = 2000$ WT experiments. The time series corresponding to these 2000 experiments are selected as the source domain data.

Using the above target domain and source domain data, the WTs in the WF can be clustered using MVIT-FCM.

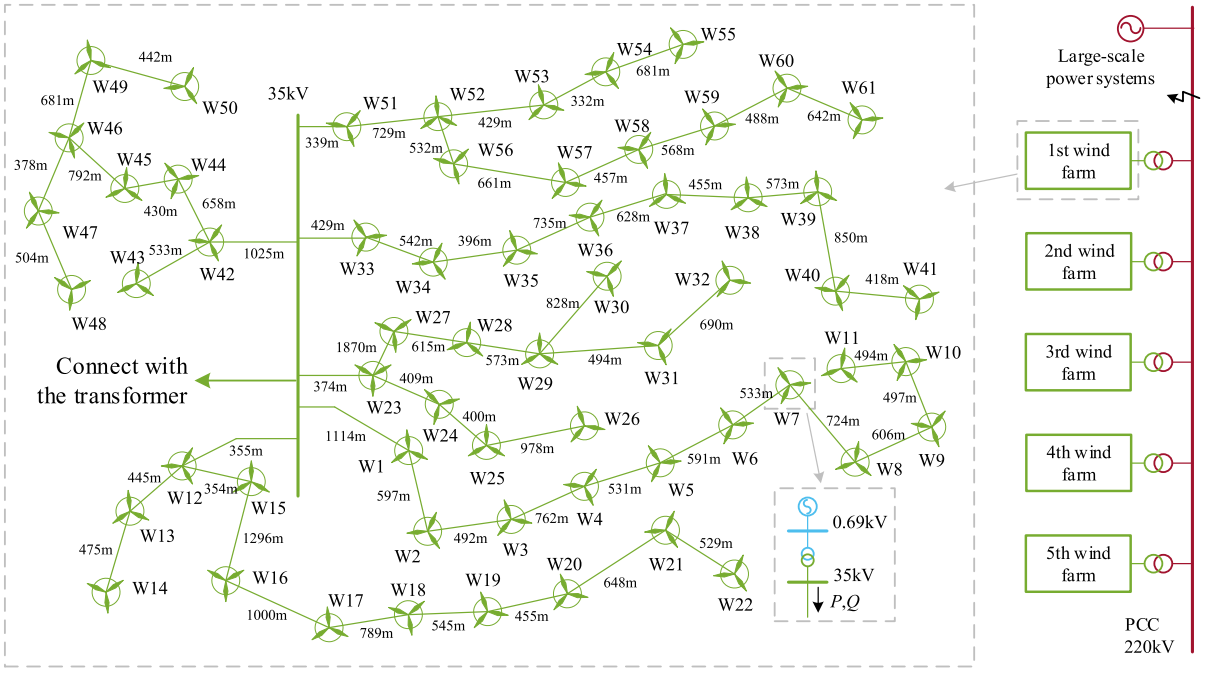


FIGURE 7. The layout of the modified wind farm consisting of 61 wind turbines.

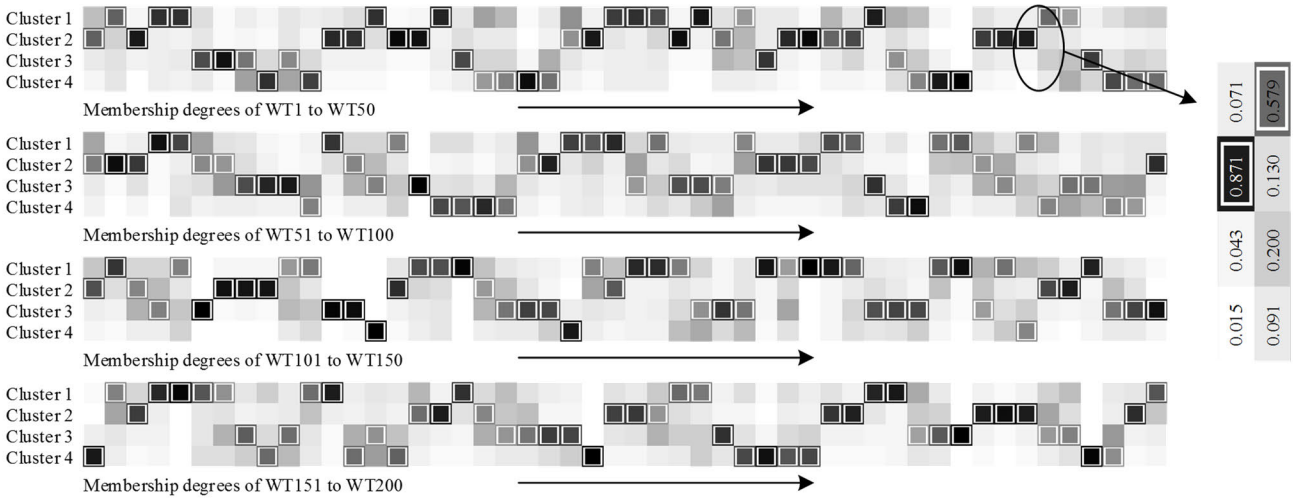


FIGURE 8. The final clustering result.

The parameters setting of MVIT-FCM is: $N = 200$, $C = 4$, $m = 2$, $\lambda = 1$, $\gamma = 0.8$, $N_l = 50$, $M = 4$, $\overline{N}_l = 50$ and $\overline{M} = 40$. In Fig. 2, we can see that not the source domain samples are directly transferred to target domain, but the clustering center $\overline{\boldsymbol{V}}_k$ and membership degree $\overline{\boldsymbol{U}}_k$ are transferred. Thus, the $\overline{\boldsymbol{V}}_k$ and $\overline{\boldsymbol{U}}_k$ are calculated based on the source domain samples using the 1st-10th columns of pseudocode listed in Fig. 2. After acquiring $\overline{\boldsymbol{V}}_k$ and $\overline{\boldsymbol{U}}_k$, these two terms are transferred and used to guide the clustering of the target domain samples. This target domain data clustering process can be achieved using the 11st-20th columns of pseudocode listed in Fig. 2, and we can get $\boldsymbol{V}_k$ and $\boldsymbol{U}_k$. Finally, the $\hat{\boldsymbol{U}}$ is calculated using (2), and the final clustering result is shown in Fig. 8. In Fig. 8, the darker of the color, the larger of the value. The largest value in each column is circled, and for each row, all the circled WTs are clustered to a group.

B. KEY PARAMETERS SELECTION AND OPTIMIZATION RESULTS

The sampling times is set to 20 for each parameter in WT. Besides, the threshold value ε_1 (ε_T) are set to 2.2 (0.1), 1.1 (0.05), 2.5 (0.5) and 0.6 (0.3) for active power, reactive power, voltage and current. The S_i and S_{Ti} of y_1-y_{56} are shown in Fig. 9 and Fig. 10, and the parameters that are larger

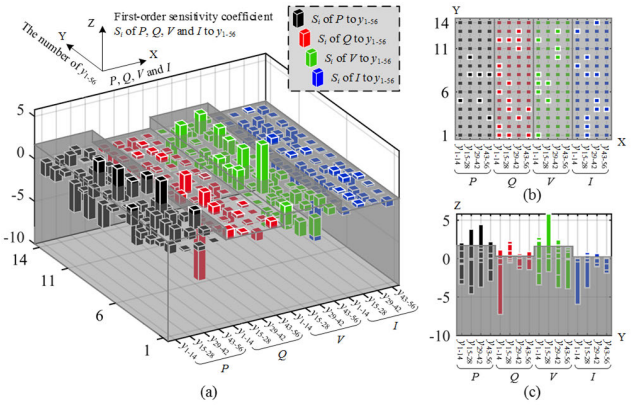


FIGURE 9. The first-order sensitivity coefficient of PCC active power, reactive power, voltage and current to y_1-y_{56}. (a) the stereogram, (b) the top view, and (c) the front view.

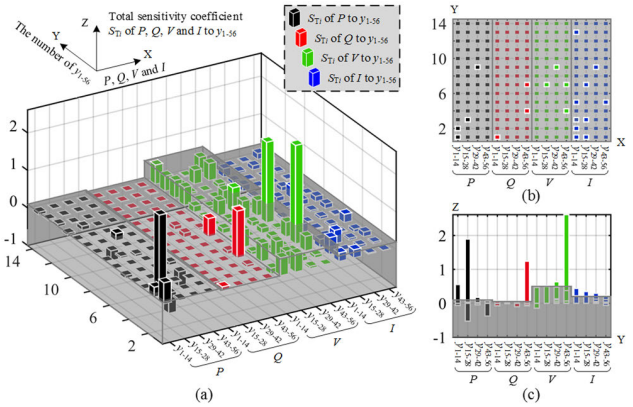


FIGURE 10. The total sensitivity coefficient of PCC active power, reactive power, voltage and current to y_1-y_{56}. (a) the stereogram, (b) the top view, and (c) the front view.

than the threshold values are selected as the key parameters. Finally, the selected key parameters are y_1, y_2, y_6, y_8, y_{12}, y_{17}, y_{19}, y_{21}, y_{22}, y_{24}, y_{32}, y_{33}, y_{37}, y_{45}, y_{46}, y_{49}.

Further, we optimize the key parameters of the equivalent WF, and the parameters setting of Q-NSGA-II is shown in TABLE 3. The convergences of f_P, f_Q, f_V and f_I are shown

TABLE 3. Parameters Setting of Q-NSGA-II.

Q-NSGA-II	Parameters	Value setting
Q-learning part	α	0.1
	γ	0.05
	ε	0.85
	QJ	6
	QI	16
	QM_i	10
	W	1
NSGA-II part	Distribution index of crossover n_c	20
	Distribution index of mutation n_m	20
	Crossover probability p_c	0.8
	Mutation probability p_m	0.1
	Population size P	60
	Iterations *Iter*	100

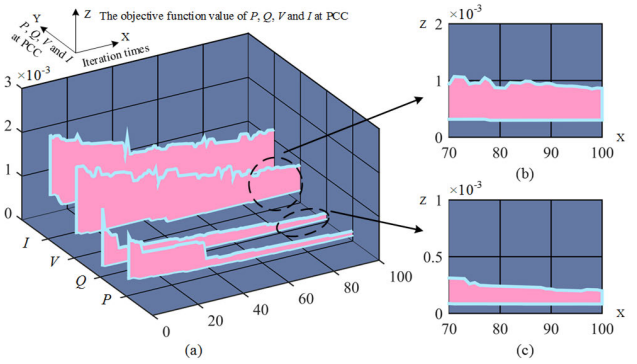


FIGURE 11. The convergence of the f_P, f_Q, f_V and f_I sub-objective with the iteration times increase. (a) the stereogram, (b) part of f_V, and (c) part of f_Q.

in Fig. 11. From Fig. 11, all the sub-objectives have converged when the iteration is around 80. Therefore, we select the final optimization results of the key parameters from the Pareto optimal solutions in 100th iteration. In this article, the fuzzy set theory [21] is adopted to determine the final solution from the Pareto optimal solutions. The final optimization results of the first equivalent WT (corresponding to Cluster 1 in Fig. 8) is shown in Fig. 12.

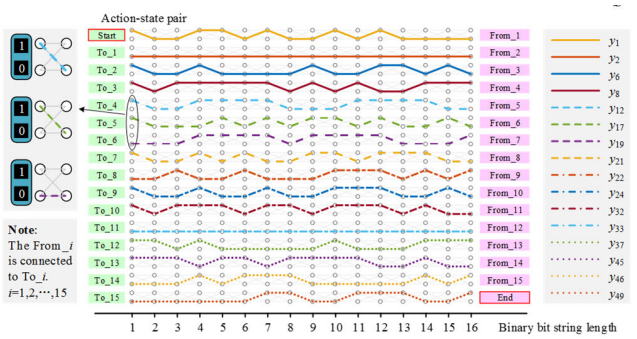


FIGURE 12. The final optimization results of key parameters in WT_{eq1}.

C. ACCURACY OF THE WFs EQUIVALENT METHODS

In this article, the CI selects multi-view time series, which provids more views to ensure the equivalent accuracies in more aspects. Further, MVIT-FCM is proposed for increasing the algorithm usability when processing large-scale WFs and increasing the stability of the clustering results. Finally, Q-NSGA-II is proposed to optimize the equivalent parameters. To illustrate the superiority of the proposed method, the proposed model (M0) is compared with several multi-machines equivalent models (M1 to M5). The descriptions of these models are listed in TABLE 4.

To verify the accuracy of the WFs equivalent models, root mean square error (e_{RMSE}), mean absolute error (e_{MAE}), and mean absolute percentage error (e_{MAPE}) are introduced in the article, which are widely used to evaluate the error of time series. The expressions of e_{RMSE}, e_{MAE} and e_{MAPE} are shown

TABLE 4. The used CI, clustering algorithm and equivalent parameters calculation method of the models.

Model	Literature	Used clustering indicator	Used clustering algorithm	Equivalent parameters calculation method
M0	This paper	Time series of active power, reactive power, voltage and current	The proposed MVIT-FCM	Q-NSGA-II based parameters optimization method
M1	[12]	The extracted features of active power, reactive power, voltage and current using auto-encoder	The proposed MVT-FCM	Capacity weighted method
M2	[12]	Time series of active power, reactive power, voltage and current	The proposed MVT-FCM	Capacity weighted method
M3	[11]	Time series of active power	Geometric template matching and quality threshold clustering	Capacity weighted method
M4	[23]	Three cloud model characteristic quantities of dynamic voltage	k-means clustering	Capacity weighted method
M5	[24]	Rotating speed of WT when fault is cleared	k-means clustering	Capacity weighted method

in (11).

$$\begin{cases} e_{\mathrm{RMSE}} = \sqrt{\sum_{l=1}^{L} (X(l) - \hat{X}(l))^2 / L} \\ e_{\mathrm{MAE}} = \sum_{l=1}^{L} \left| X(l) - \hat{X}(l) \right| / L \\ e_{\mathrm{MAPE}} = \sum_{l=1}^{L} \left| (X(l) - \hat{X}(l)) / X(l) \right| / L \end{cases} \tag{11}$$

The e_{RMSE}, e_{MAE} and e_{MAPE} of M0-M5 are shown in TABLE 5. In TABLE 5, by comparing the value of e_{RMSE}, e_{MAE} and e_{MAPE} in M0-M5, M0 performs the same well or even better in the accuracy aspect when comparing with other five models. The reasons are listed as follows.

Firstly, as for M0 and M1/M2, we can see that these three models use multi-view CI and multi-view clustering algorithms, and acquire accurate results in active power, reactive power, voltage and current equivalences. Yet, the proposed model performs even better. This is because the proposed model acquires better equivalent parameters using Q-NSGA-II, and this parameters optimization process is significantly helpful in increasing the equivalent accuracy. Although the Q-NSGA-II based optimization method is slightly more complex and time-consuming than the traditional capacity weighted method in the equivalent parameters calculation process, it only needs to be carried out once and the calculation results can be reused. Thus, the slightly more tedious process is worthy because better equivalent accuracy is acquired.

Secondly, with M0 and M3/M4/M5, the compared models all use single-view CI for WTs clustering, thus, their accuracies are higher in the aspect corresponding to their CIs. Specifically, M3 uses active power as the CI, so its active power accuracy is higher than other physical quantities. Similarly, for the M4 model, the voltage accuracy is higher, this is because the dynamic voltage is selected as the CI in this model. M5 clusters the WTs using their rotating speed, which is a physical quantity relevant to active power, so this model performs better in the equivalence of active power. Conversely, the proposed model uses multi-view CI, thus, its equivalent accuracies are ensured in more aspects. In addition, the compared models use traditional capacity weighted method for equivalent parameters calculation, and the results are not the same well when comparing with Q-NSGA-II based optimization method.

TABLE 5. The simulation errors (E_{RMSE}, E_{MAE} and E_{MAPE}) in different models.

Models	Physical quantities	Error types		
		e_{RMSE}	e_{MAE}	e_{MAPE}
M0 (Proposed model)	Active power	7.24×10^{-2}	1.17×10^{-2}	1.33×10^{-2}
	Reactive power	4.73×10^{-2}	2.71×10^{-2}	3.04×10^{-2}
	Voltage	6.77×10^{-3}	1.54×10^{-3}	1.89×10^{-3}
	Current	4.35×10^{-3}	3.57×10^{-3}	4.00×10^{-3}
M1	Active power	9.04×10^{-2}	1.44×10^{-2}	1.67×10^{-2}
	Reactive power	5.47×10^{-2}	2.92×10^{-2}	3.60×10^{-2}
	Voltage	8.74×10^{-3}	1.96×10^{-3}	2.37×10^{-3}
	Current	5.56×10^{-3}	4.43×10^{-3}	5.18×10^{-3}
M2	Active power	1.45×10^{-1}	2.47×10^{-2}	2.76×10^{-2}
	Reactive power	2.23×10^{-1}	2.39×10^{-1}	2.67×10^{-1}
	Voltage	1.65×10^{-2}	4.36×10^{-3}	5.34×10^{-3}
	Current	8.91×10^{-3}	6.58×10^{-3}	7.34×10^{-3}
M3	Active power	1.76×10^{-1}	2.95×10^{-2}	3.29×10^{-2}
	Reactive power	2.35×10^{-1}	2.78×10^{-1}	3.28×10^{-1}
	Voltage	2.57×10^{-2}	5.52×10^{-3}	6.27×10^{-3}
	Current	1.14×10^{-2}	7.53×10^{-3}	9.25×10^{-3}
M4	Active power	1.45×10^{-1}	3.18×10^{-2}	3.77×10^{-2}
	Reactive power	3.05×10^{-1}	3.28×10^{-1}	3.68×10^{-1}
	Voltage	2.22×10^{-2}	5.91×10^{-3}	6.72×10^{-3}
	Current	1.17×10^{-2}	8.96×10^{-3}	1.00×10^{-2}
M5	Active power	1.86×10^{-1}	4.09×10^{-2}	4.73×10^{-2}
	Reactive power	3.78×10^{-1}	3.61×10^{-1}	4.49×10^{-1}
	Voltage	2.76×10^{-2}	7.31×10^{-3}	8.26×10^{-3}
	Current	1.47×10^{-2}	1.11×10^{-2}	1.24×10^{-2}

D. PERFORMANCE OF MVIT-FCM WHEN PROCESSING DIFFERENT SCALE WFs

Through the analysis in Section VI-C, M0, M1and M2 acquire satisfactory performances, and it seems that the proposed MVIT-FCM is not much helpful when comparing with

MVT-FCM. This is because the scale of the test WFs system is not large enough, and MVT-FCM is still valid. Yet, if the scale of the WFs is quite huge, the data-size of the multi-view time series-based CI would be extremely huge and might exceed the memory of the computer. Thus, the effectiveness of MVIT-FCM and MVT-FCM when processing different scales WFs needs to be tested.

In order to verify the calculation ability when processing different scale WFs, we test the performances of MVIT-FCM and MVT-FCM under different number of clustering samples, i.e., different number of WTs. In Fig. 7, there are totally 200WTs, and we randomly copy different number of samples to constitute the new dataset to be clustered. The computer configuration is Intel(R) Core(TM) i5-8250, CPU@1.60 GHz, 8.00 GB of RAM, and Fig. 13 shows the maximum CPU utilization (MCU) of MVIT-FCM and MVT-FCM under different numbers of samples. It should be noted that the MCU for each number of samples are calculated after 100 repeated runs. Thus, Fig. 13 shows the expectation and variance of the MCU.

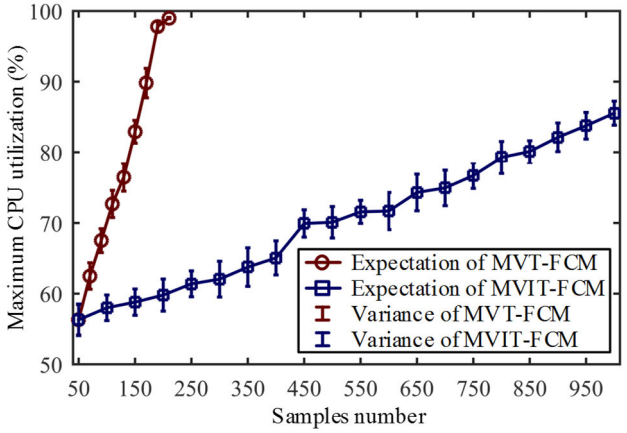


FIGURE 13. Maximum CPU utilization of MVIT-FCM and MVT-FCM.

It can be seen that when the number of data samples is 200, the MCU of MVT-FCM is close to 100%. In other words, MVT-FCM can only deal with the WFs equivalent problem with about 200 WTs. However, with MVIT-FCM, the MCU increases slowly with the increasing of samples number. The computer with the author's personal computer can handle the clustering problem with more than 1000 WTs. Thus, MVIT-FCM has the ability of handling the super-large scale of WFs equivalent problem with any number of WTs.

Further, in the proposed MVIT-FCM, the dataset is divided into several blocks. Then, MVT-FCM is applied to each block, and different blocks are coupled with weight coefficient and updating clustering center. To test whether the partition mechanism would influence the clustering accuracy, we test the performance of MVIT-FCM and MVT-FCM. In addition, the traditional MV-FCM is also used for comparison. Three popular validity indices are adopted in this article, i.e., DVI, SI, and DBI [25], for verifying clustering performance of these algorithms. Larger values of DVI and SI indicate better clustering performance. In contrast, smaller values of DBI are preferred. We calculate the DVI, SI, and DBI of the three clustering algorithms using different number of data samples under 100 repeated runs, and these values are listed in TABLE 6.

TABLE 6. Values of DVI, SI, and DBI in MVIT-FCM and MVT-FCM.

Samples number	Algorithm	Value	Metrics		
			DVI	SI	DBI
50	MVIT-FCM	Mean	0.9244	0.9312	0.5548
		Std	0	0	0
	MVT-FCM	Mean	0.9244	0.9312	0.5548
		Std	0	0	0
	MV-FCM	Mean	0.8567	0.8847	0.7725
		Std	5.18×10^{-5}	2.36×10^{-5}	6.52×10^{-5}
100	MVIT-FCM	Mean	0.9089	0.9144	0.5678
		Std	0	0	0
	MVT-FCM	Mean	0.9089	0.9144	0.5678
		Std	0	0	0
	MV-FCM	Mean	0.8664	0.8758	0.7425
		Std	3.44×10^{-5}	4.85×10^{-5}	5.26×10^{-5}
150	MVIT-FCM	Mean	0.9381	0.9462	0.5468
		Std	0	0	0
	MVT-FCM	Mean	0.9381	0.9462	0.5468
		Std	0	0	0
	MV-FCM	Mean	0.8134	0.8156	0.7889
		Std	2.67×10^{-5}	1.61×10^{-5}	7.64×10^{-6}
200	MVIT-FCM	Mean	0.9244	0.9301	0.5540
		Std	0	0	0
	MVT-FCM	Mean	0.9244	0.9301	0.5540
		Std	0	0	0
	MV-FCM	Mean	0.8804	0.8544	0.6213
		Std	4.86×10^{-5}	3.44×10^{-5}	5.21×10^{-5}
250	MVIT-FCM	Mean	0.9331	0.9284	0.5347
		Std	0	0	0
	MVT-FCM	Mean	0.9331	0.9284	0.5347
		Std	0	0	0
	MV-FCM	Mean	0.8742	0.8661	0.6815
		Std	3.48×10^{-5}	2.48×10^{-5}	3.02×10^{-5}
300	MVIT-FCM	Mean	0.9422	0.9308	0.5110
		Std	0	0	0
	MVT-FCM	Mean	0.9422	0.9308	0.5110
		Std	0	0	0
	MV-FCM	Mean	0.8679	0.8749	0.6441
		Std	4.33×10^{-5}	2.57×10^{-5}	2.68×10^{-5}

With MV-FCM, its mean values of DVI and SI in MVT-FCM are apparently smaller than MVIT-FCM and MVT-FCM, and the DBI values of MV-FCM are larger. Thus, the clustering accuracy of MV-FCM is not well-performed. Besides, the standard deviation values in MV-FCM are larger than 0, which means different clustering results are acquired in the 100 repeated runs. This is because the MV-FCM without the introducing of transfer learning is sensitive to the initialization, resulting in different clustering results. Thus, the applying of transfer learning mechanism is quite helpful to increase the stability of clustering results.

With the comparison of MVIT-FCM and MVT-FCM, the mean values and the standard deviation values of DVI, SI and DBI in both algorithms are totally the same. This indicates that the same clustering results are acquired in both algorithms. In addition, the standard deviation values

of DVI, SI, and DBI in both algorithms are 0. This indicates that the same clustering results are acquired in the 100 repeated runs. Thus, MVIT-FCM is the same well-performed when comparing with MVT-FCM, and the partition mechanism would not influence the clustering accuracy or performance of MVIT-FCM, proving that MVIT-FCM is an powerful tool in processing large-scale and huge-dataset without losing any clustering accuracy.

E. SUPERIORITY OF THE EQUIVALENT PARAMETERS CALCULATION METHOD

In order to illustrate the superiority of Q-NSGA-II in optimizing the key parameters, we design several parameters optimization manners, which are shown TABLE 7.

The sub-objective function fitness values of different parameters optimization manners under different iterations are calculated, as shown in Fig. 14. It should be noted that since Q-NSGA-II has converged when iteration is 100, so the fitness values under 100-th iteration are taken as the Pareto frontier. Besides, there are 4 sub-objective functions in key parameters optimization, for better visualization, we put f_P and f_Q in the same figure, and put f_V and f_I in the same picture. From the figure, we can find that the proposed manner (Q-NSGA-II optimizing key parameters) always leads the optimization direction. In other words, its optimization efficiency and speed are better than other manners. The reason can be summarized as two aspects. On the one hand, the Sobol'-based criterion for selecting key parameters significantly decreases the number of the parameters to be optimized, reducing the heavy burdens for the optimization algorithms. On the other hand, the Q-NSGA-II algorithm introduces the Q-learning technique, increasing the intelligent degree of the algorithm and making the optimization process more efficient. Thus, the proposed manner is not only helpful for increasing the equivalent accuracy, but also has the advantage of efficient solving speed.

TABLE 7. The designed parameters optimization manners.

Manner	Optimization method	Optimized parameters
1	Q-NSGA-II	Key parameters, i.e., $\boldsymbol{y}_{\text{key}}$
2	Q-NSGA-II	All parameters, i.e., y_1-y_{56}
3	NSGA-II	Key parameters, i.e., $\boldsymbol{y}_{\text{key}}$
4	NSGA-II	All parameters, i.e., y_1-y_{56}

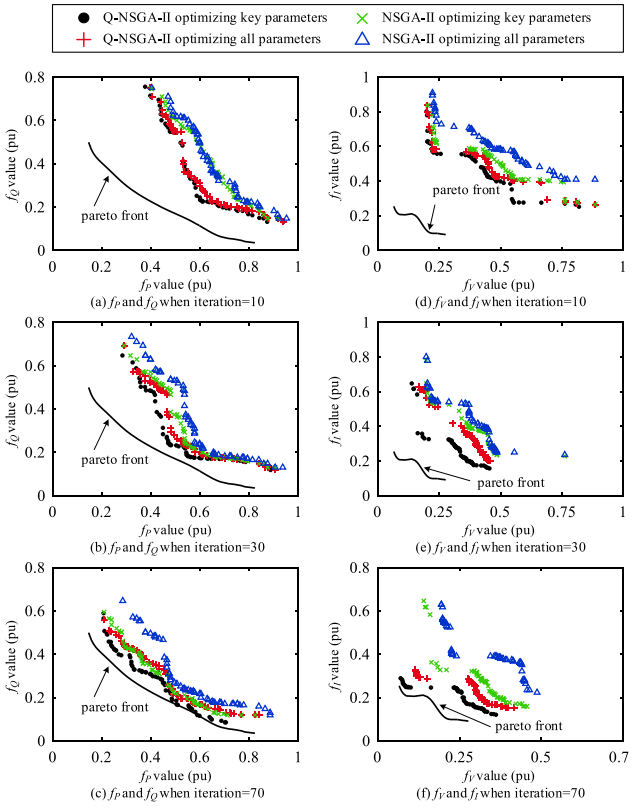


FIGURE 14. Fitness value of subobjective function under different iterations. (a) fitness value of f_P and f_Q when iteration = 10, (b) fitness value of f_P and f_Q when iteration = 30, (c) fitness value of f_P and f_Q when iteration = 70, (d) fitness value of f_V and f_I when iteration = 10, (e) fitness value of f_V and f_I when iteration = 30, and (f) fitness value of f_V and f_I when iteration = 70.

VII. CONCLUSION

This article proposes an equivalent method for large-scale WFs using incremental clustering and key parameters optimization. Based on the simulation study of the modified WFs systems, several conclusions can be drawn:

(1) Through the multi-view consideration in CI selection and the key parameters optimization process, the equivalent accuracy of the proposed model is higher when comparing with other state-of-the-art models.

(2) Through the introducing of incremental technique into MVIT-FCM, the proposed method can deal with the equivalent problems of large-scale WFs without requiring to consider the computer configuration.

(3) With the help of key parameters selection and Q-learning, Q-NSGA-II performs efficient and helps increase the equivalent accuracy, proving the proposed method a powerful tool in accurate WFs dynamic simulation.

APPENDIX

A. PARAMETERS OF THE EQUIVALENT UNIT

See Table 8.

B. VOLTAGE DIP CALCULATION IN EACH WIND TURBINE

In the regulation, when there is three-phase symmetrical voltage swell, the reactive current injected into the WT from the power system should meet the following conditions.

$$I_T \geq 1.5 \times (0.9 - V_N) \tag{A1}$$

where V_n denotes the voltage in the n-th WT port.

In order to relieve LVRT pressure of the WT, the greater than or equal in (A7) is set to equal, and the voltage dip calculation method in each WT is shown as follows.

The structure of the collector line connecting the WTs, and the vector diagram between 2 nodes are shown as follows.

TABLE 8. Parameters of the equivalent unit.

Component	Parameters
Wind turbine	***Generator:***
	Nominal power (y_1); L-L voltage (y_2);
	Stator inductance (y_3); Stator resistance (y_4);
	Rotor inductance (y_5); Rotor resistance (y_6);
	Magnetizing inductance (y_7);
	Inertia constant (y_8); Friction factor (y_9).
	Converters:
	Grid-side converter maximum current (y_{10}); Grid-side coupling inductor (y_{11}, y_{12});
	Nominal DC bus voltage (y_{13});
	DC bus capacitor (y_{14}); Line filter capacitor (y_{15}).
	Turbine:
	Nominal mechanical output power (y_{16});
	Wind speed at nominal speed (y_{17}).
	Drive train:
	Wind turbine inertia constant (y_{18});
	Shaft spring constant refer to high-speed shaft (y_{19});
	Shaft mutual damping (y_{20}).
	Control system:
	DC bus voltage regulator gains (y_{21}, y_{22});
	Grid-side converter current regulator gains (y_{23}, y_{24});
	Speed regulator gains (y_{25}, y_{26});
	Rotor-side converter current regulator gains (y_{27}, y_{28});
	Q and V regulator gains (y_{29}, y_{30});
	Pitch controller and compensation gain (y_{31}, y_{32}, y_{33});
	Maximum pitch angle (y_{34});
	Maximum rate of change of pitch angle (y_{35}).
Box transformer (25kV/575V)	Nominal power (y_{36});
	Winding 1 parameters:
	L-L voltage, inductance and resistance (y_{37}, y_{38}, y_{39});
	Winding 2 parameters:
	L-L voltage, inductance and resistance (y_{40}, y_{41}, y_{42});
	Magnetization inductance and resistance (y_{43}, y_{44}).
Main transformer (120kV/25kV)	Nominal power (y_{45});
	Winding 1 parameters:
	L-L voltage, inductance and resistance (y_{46}, y_{47}, y_{48});
	Winding 2 parameters:
	L-L voltage, inductance and resistance (y_{49}, y_{50}, y_{51});
	Magnetization inductance and resistance (y_{52}, y_{53}).
Collector network	Positive-sequence resistances (y_{54});
	Positive-sequence inductances (y_{55});
	Positive-sequence capacitances (y_{56}).

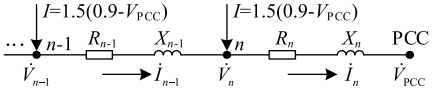


FIGURE 15. Structure of branch in the wind farm.

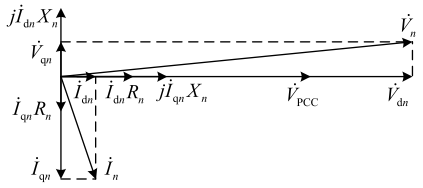


FIGURE 16. Vector diagram of the branch between node n and PCC.

According to Fig. 15 and Kirchhoff current law, the expression of I_{qn} is

$$I_{qn} = 1.5n(0.9 - V_{PCC}) \tag{A2}$$

Then, through the vector relationship between the voltage and current of in Fig. 16, it is easy to obtain

$$V_{qn} = I_{dn}X_n - I_{qn}R_n \tag{A3}$$

$$V_{dn} - V_{PCC} = I_{qn}X_n + I_{dn}R_n \tag{A4}$$

Besides, in the power system short circuit calculation, the phase angle difference between different nodes is ignored, generally. In other words, the phase angle of $\dot{V}_{n+1}$ and $\dot{V}_n$ are totally the same, and we add the additional condition

$$V_{qn} = 0 \tag{A5}$$

Therefore, $V_n = V_{dn}$.

We rewrite the equation (A2) to (A5) considering $V_n = V_{dn}$, which is as follows

$$\begin{cases} I_{qn} = 1.5n(0.9 - V_{PCC}) \\ I_{dn}X_n - I_{qn}R_n = 0 \\ V_n - V_{PCC} = I_{qn}X_n + I_{dn}R_n \end{cases} \tag{A6}$$

Overall, we can write 3 equations for each node. If there are N nodes in this branch, the number of the equations would be $3N$. In these $3N$ equations, the unknown quantities include V_i, I_{di}, I_{qi} $(i = 1, 2, \ldots, N)$ and V_{N+1}, totally $3N + 1$, where V_{N+1} is equal to V_{PCC}, which is a known quantity. Therefore, the $3N$ equations contain $3N$ unknown quantities, and can be solved.

REFERENCES

[1] *The 13th Five-Year Plan for Power Development*, National Development and Reform Commission, Beijing, China, 2016.

[2] Y. Zhou, L. Zhao, and W.-J. Lee, "Robustness analysis of dynamic equivalent model of DFIG wind farm for stability study," *IEEE Trans. Ind. Appl.*, vol. 54, no. 6, pp. 5682–5690, Nov. 2018.

[3] L. M. Fernández, C. A. García, J. R. Saenz, and F. Jurado, "Equivalent models of wind farms by using aggregated wind turbines and equivalent winds," *Energy Convers. Manage.*, vol. 50, no. 3, pp. 691–704, Mar. 2009.

[4] R. Fang, R. Shang, M. Wu, C. Peng, and X. Guo, "Application of gray relational analysis to k-means clustering for dynamic equivalent modeling of wind farm," *Int. J. Hydrogen Energy*, vol. 42, no. 31, pp. 20154–20163, Aug. 2017.

[5] J. Zou, C. Peng, H. Xu, and Y. Yan, "A fuzzy clustering algorithm-based dynamic equivalent modeling method for wind farm with DFIG," *IEEE Trans. Energy Convers.*, vol. 30, no. 4, pp. 1329–1337, Dec. 2015.

[6] P. Chao, W. Li, X. Jin, J. Qi, and X. Li, "An active power response based practical equivalent method for DFIG wind farms," *Proc. CSEE*, vol. 38, no. 6, pp. 1639–1646, Mar. 2018.

[7] L. Lin, X. Pan, L. Zhang, and S. Zhao, "The K-means clustering algorithm for wind farm based on immune-outlier data and immune-sensitive initial center," *Proc. CSEE*, vol. 36, no. 20, pp. 5461–5468, Oct. 2016.

[8] P. Wang, Z. Zhang, Q. Huang, N. Wang, X. Zhang, and W.-J. Lee, "Improved wind farm aggregated modeling method for large-scale power system stability studies," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6332–6342, Nov. 2018.

[9] W. Li, P. Chao, X. Liang, J. Ma, D. Xu, and X. Jin, "A practical equivalent method for DFIG wind farms," *IEEE Trans. Sustain. Energy*, vol. 9, no. 2, pp. 610–620, Apr. 2018.

[10] X. Zhang, L. Li, X. Hu, X. Wan, and X. Zhou, "Wind farm dynamic equivalence based on clustering by output time series data of wind turbine generators," *Power Syst. Technol.*, vol. 39, no. 10, pp. 2787–2793, Oct. 2015.

[11] J. Han, S. Miao, Y. Li, W. Yang, and H. Yin, "A wind farm equivalent method based on multi-view transfer clustering and stack sparse auto encoder," *IEEE Access*, vol. 8, pp. 92827–92841, 2020.

[12] J. Han, S. Miao, Y. Li, W. Yang, and T. Zheng, "A multi-view and multi-scale transfer learning based wind farm equivalent method," *Int. J. Electr. Power Energy Syst.*, vol. 117, May 2020, Art. no. 105740.

[13] F. Gonzalez-Longatt, P. Regulski, P. Wall, and V. Terzija, "Procedure for estimation of equivalent model parameters for a wind farm using post-disturbance on-line measurement data," in *Proc. 2nd IEEE PES Int. Conf. Exhib. Innov. Smart Grid Technol.*, Dec. 2011, pp. 1–7.

[14] T. C. Havens, J. C. Bezdek, C. Leckie, L. O. Hall, and M. Palaniswami, "Fuzzy c-Means algorithms for very large data," *IEEE Trans. Fuzzy Syst.*, vol. 20, no. 6, pp. 1130–1146, Dec. 2012.

[15] I. M. Sobol, "Sensitivity estimates for nonlinear mathematical models," *Matem Mod*, vol. 2, no. 1, pp. 112–118, 1993.
[16] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, Apr. 2002.
[17] X. Zhang, T. Bao, T. Yu, B. Yang, and C. Han, "Deep transfer Q-learning with virtual leader-follower for supply-demand stackelberg game of smart grid," *Energy*, vol. 133, pp. 348–365, Aug. 2017.
[18] Y. Jiang, F.-L. Chung, S. Wang, Z. Deng, J. Wang, and P. Qian, "Collaborative fuzzy clustering from multiple weighted views," *IEEE Trans. Cybern.*, vol. 45, no. 4, pp. 688–701, Apr. 2015.
[19] Z. Deng, Y. Jiang, F.-L. Chung, H. Ishibuchi, K.-S. Choi, and S. Wang, "Transfer prototype-based fuzzy clustering," *IEEE Trans. Fuzzy Syst.*, vol. 24, no. 5, pp. 1210–1232, Oct. 2016.
[20] S. Jialin Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.
[21] C. Peng and H. Sun, "Multi-objective optimization power dispatch based on non-dominated sorting differential evolution," *Proc. CSEE*, vol. 29, no. 34, pp. 71–76, Dec. 2009.
[22] Y. Su, Z. Mi, and Y. Wang, "Applicability and improvement of common used equivalent methods for wind farms," *Power Syst. Technol.*, vol. 34, pp. 175–180, 2010.
[23] M. Zhou, J. Ge, and G. Li, "Cloud model based DFIG wind farm dynamic voltage equivalence method," *Proc. CSEE*, vol. 35, no. 5, pp. 1097–1105, Mar. 2015.
[24] Z. Mi, X. Su, Q. Yang, Y. Wang, and T. Wu, "Multi-machine representation method for dynamic equivalent model of wind farms," *Trans. China Electrotech. Soc.*, vol. 25, no. 5, pp. 162–169, 2010.
[25] A. Fahad, N. Alshatri, Z. Tari, A. Alamri, I. Khalil, A. Y. Zomaya, S. Foufou, and A. Bouras, "A survey of clustering algorithms for big data: Taxonomy and empirical analysis," *IEEE Trans. Emerg. Topics Comput.*, vol. 2, no. 3, pp. 267–279, Sep. 2014.

JI HAN (Student Member, IEEE) was born in Liaoning, China, in 1993. He received the B.S. degree in electrical engineering from the Huazhong University of Science and Technology (HUST), Wuhan, China, in 2016, where he is currently pursuing the Ph.D. degree. His research interests include application of machine learning in power systems, renewable energy technology, and fault diagnosis of power systems.

SHIHONG MIAO (Member, IEEE) was born in Hubei, China, in 1963. He received the M.S. and Ph.D. degrees in electrical engineering from the Huazhong University of Science and Technology (HUST), Wuhan, China, in 1996 and 2003, respectively.

He is currently a Professor with the School of Electrical and Electronic Engineering, HUST. His research interests include power system protective relaying and control, renewable energy technology, novel distribution network and micro-grid technology, and compressedair energy storage (CAES) technology.

YAOWANG LI (Member, IEEE) was born in Hefei, China, in 1993. He received the B.S. and Ph.D. degrees in electrical engineering from the Huazhong University of Science and Technology (HUST), Wuhan, China, in 2015 and 2020, respectively.

He is currently a Postdoctoral Research Fellow with the Department of Electrical Engineering, Tsinghua University, Beijing, China. His research interests include multi-energy systems, cloud energy systems, adiabatic compressed air energy storage, and renewable energy modeling.

HAORAN YIN was born in Hebei, China, in 1997. He received the B.S. degree in electrical engineering from the Huazhong University of Science and Technology (HUST), Wuhan, China, in 2019, where he is currently pursuing the M.S. degree.

His research interests include application of machine learning in power systems and fault diagnosis of power systems.

DI ZHANG was born in Henan, China, in 1992. He received the B.S. degree in electrical engineering from the Huazhong University of Science and Technology (HUST), Wuhan, China, in 2015, where he is currently pursuing the Ph.D. degree.

His research interests include the operation and control of power system integrated with renewable energy.

WEICHEN YANG was born in Henan, China, in 1995. He received the B.S. degree in electrical engineering from North China Electric Power University, Baoding, China, in 2017. He is currently pursuing the Ph.D. degree with the Huazhong University of Science and Technology (HUST).

His research interests include high-voltage direct current (HVDC) technology and renewable energy technology.

QINGYU TU was born in Hubei, China, in 1994. He received the B.S. degree in electrical engineering from the Huazhong University of Science and Technology (HUST), Wuhan, China, in 2017, where he is currently pursuing the Ph.D. degree.

His research interests include renewable energy technology and the modeling of wind power generation.