energies

MDPI

Article

An Equivalent Model of Wind Farm Based on Multivariate Multi-Scale Entropy and Multi-View Clustering

Ji Han [1], Li Li [1], Huihui Song [1,*], Meng Liu [2], Zongxun Song [3] and Yanbin Qu [1]

[1] School of New Energy, Harbin Institute of Technology at Weihai, Weihai 264200, China
[2] Weihai Power Supply Company, State Grid Shandong Electric Power, Ltd., Jinan 264200, China
[3] Electric Power Research Institute, State Grid Shandong Electric Power, Ltd., Weihai 250003, China
* Correspondence: songhh@hitwh.edu.cn; Tel.: +86-156-6236-2246

Abstract: Wind farm (WF) equivalence is an effective method to achieve accurate and efficient simulation of large-scale WF. Existing equivalent models are generally suitable for one certain or very few scenarios, and have difficulty reflecting the multiple aspects of dynamic processes of WF. Aiming at these problems, this paper proposes an equivalent model of WF based on multivariate multi-scale entropy (MMSE) and multi-view clustering. Firstly, the influence of the factors on the dynamic process of the wind turbine (WT) is discussed, including control mode, wind speed and its wake effect, resistance of crowbar resistor and so on. The relationship between these factors and the dynamic equivalence of WF is analyzed. Secondly, an overview of MMSE is given, and the applicability of MMSE on WF equivalence is analyzed. On this basis, this paper proposes the extraction process of a WT clustering indicator using MMSE. Then, the multi-view fuzzy C means (MV-FCM) algorithm is used for the clustering of WTs, and the equivalent model of WF is obtained after calculating the equivalent parameters. Finally, the IEEE14 power system including WF is simulated. The results show that the equivalent model could be applied to dynamic process simulation in various fault scenarios of power systems, and the error is small when the cluster number is 4. Compared with the detailed model, the simulation time of the WF equivalent model proposed in this paper is shortened by 86%, and the simulation accuracy is improved by about 44% compared with the comparative model.

Keywords: wind farm; multivariate multi-scale entropy; WF equivalence; data mining; cluster; multi-scenario

Citation: Han, J.; Li, L.; Song, H.; Liu, M.; Song, Z.; Qu, Y. An Equivalent Model of Wind Farm Based on Multivariate Multi-Scale Entropy and Multi-View Clustering. *Energies* **2022**, *15*, 6054. https://doi.org/10.3390/en15166054

Academic Editor: Robin Roche

Received: 24 June 2022
Accepted: 18 August 2022
Published: 21 August 2022

1. Introduction

According to the wind power report released by the Global Wind Energy Council (GWEC) [1] in 2021, the global new wind power installed capacity is 93.6 GW, and the cumulative installed capacity reaches 837 GW, with a year-on-year increase of 12.4%. The proportion of wind power in power demand continues to increase, and new installations in China, the United States, Brazil, Vietnam and the United Kingdom account for 75.1% of the global total. With the rapid increase of the scale of wind power grid connection, the flexibility and security stability of power systems are affected and impacted. Therefore, the requirement for analysis and calculation of large-scale wind power integration into power systems is increasing [2,3]. WF often has dozens or even hundreds of WTs. If each WT is modeled separately, it will greatly increase the complexity and simulation calculation time of the power system simulation model, and even face the problem of "dimension disaster" [4]. On the basis of satisfying the simulation accuracy, the equivalent modeling of WF provides a new idea for solving this problem.

In [5], the WF was equivalent to a WT, and its capacity was equal to the sum of the capacities of all WTs. However, the wind speed distribution in large WF is uneven, and the WTs are at different operating points. So, this kind of WF equivalent model usually has a large error. Therefore, in recent years, the research on the equivalent value of WF mainly

focuses on the multi-machine equivalent value, which mainly includes the construction of WF equivalence grouping indicators [6–9], the optimization of clustering algorithm in the equivalence process, and the new equivalence method suitable for multi-scenario analysis [10–12]. At present, the indicators used for WF equivalence are generally selected as wind speed [6], action behavior of pitch angle [7], input of crowbar protection [8], internal state quantity of WTs [9], active power output of WT, etc. This kind of research is relatively mature now. In terms of clustering algorithms, traditional clustering algorithms, such as K-means algorithm and fuzzy C-means clustering algorithm, have been widely and successfully applied to the problem of WF equivalence. Reference [13] used a K-means algorithm to cluster the original wind speed matrix and the matrix composed of 13 state variables, respectively, to the equivalence of the WF. In order to improve the accuracy and efficiency of the equivalence algorithm, the fuzzy clustering algorithm was introduced into the multi-machine equivalence of WF in [14]. However, with the increase of the number of WTs, the amount of calculation will increase. In order to avoid the complex calculation process of clustering before determining the number of clusters in traditional fuzzy clustering and ensure the practicability of the clustering results, reference [15] used the fuzzy C-means clustering algorithm for WF modeling. The algorithm is simple and has fast convergence. In addition, reference [16] used the support vector machine (SVM) algorithm to form a set of clusters for each wind speed and direction. The wind speed and direction data of the WF determine the occurrence probability of each group of data. Based on the groups with the highest probability and the number of clusters, the clustering model for the year is determined by probability clustering method. However, the distribution of wind speed and wind direction in actual WF is uncertain, and it is not suitable for the cluster classification of WF throughout the year. Nowadays, the research on the clustering algorithm in WF equivalence mainly focuses on how to improve clustering accuracy. Although favorable results have been acquired through the above studies, limitations still exist. Specifically, the equivalent model is generally applicable to a relatively single scenario. In other words, with the change of the operating conditions in the WF, the clustering results of the WTs would also be changeable, and thus the applicability of the equivalent model is greatly limited.

To solve the above problems, reference [10] considered the wake effect and the influence of low voltage ride-through characteristics of WTs on the response capability of WF, and proposed a three-step aggregation method for WF, which improves the simulation efficiency of the WF equivalent model in low voltage cases. Reference [11] used time series to cluster WTs, which ensures that the clustering results of WF are effective within a period of time. Apparently, this method can effectively reduce the number of WTs clustering. However, this method only uses the active power in the WTs clustering, so the accuracies for other variables except for active power cannot be guaranteed. In recent years, the big data technology [17] and artificial intelligence technology [18] have been well applied in data mining and feature extraction in data. Reference [12] applied big data mining technology to WF equivalence, and proposed a WF equivalence method based on a cloud model. The established equivalent model can be used in various types of interlocking off-line analysis of WTs. However, the model takes the voltage of the WT in a single time point as the object of feature extraction, and it can only reflect the characteristics of the WT at one moment, and cannot reflect the dynamic characteristics of the output of the WT in a period of time. Therefore, this method is also limited in the analysis of the dynamic characteristics of the WF output.

In this paper, a WF equivalent model based on MMSE is proposed. The model takes the time series of the WT port state as the mining object, and obtains the operating characteristics of the WT by extracting its MMSE. By using the advantages of MMSE in describing uncertain fuzzy systems and overcoming the difficulties in processing time series data with classical entropy theory, WT clustering and WF equivalence are completed. It can effectively solve the problem that the current WF equivalent model has a single applicable scene and is difficult to effectively reflect the dynamic characteristics of the WF output. At

the same time, in order to verify the applicability of the equivalent WF model in various power system fault situations, that is, the multi-scenario output dynamic characteristics of the equivalent WF, this paper sets up various power system fault conditions such as wind speed, wind direction and fault location. The example analysis shows that the equivalent WF model proposed in this paper has the following advantages: it can more accurately reflect the overall operating characteristics of the WF in the case of multi-scenario power system faults, and the simulation time is shortened by 86% compared with the detailed model, the accuracy is increased by 44% compared with the comparison model.

The remainder of this paper is organized as follows. Section 2 analyses the key factors influencing the dynamic characteristics of WT output and their relationship with the equivalent WF model. In Section 3, we propose an extraction method of the WT clustering index based on MMSE. In Section 4, we use the MV-FCM WT clustering method for equivalent model of WF. Simulations are provided in Section 5 to validate the approach. We conclude the paper in Section 6.

2. Key Factors Affecting the Dynamic Characteristics of WT Output

In this paper, the output dynamic characteristics of the WT refers to the change process of the WT active and reactive power outputs during the short-circuit fault in the external power grid of the WT. The output dynamic characteristics of WTs are affected by multiple factors. This chapter analyzes the key factors that affect the output dynamic characteristics of WTs and their relationship with the equivalence of WF.

2.1. Analysis of Key Factors

There are many factors that influence the dynamic characteristics of the WT output, such as the transition of the WT in fault type, the effect of wind speed and its weak effect, and the direction of wind affect the active power output of the WT. The pry bar protection resistance can change the reactive power dynamic of the WT during the fault, and the position of the fault will affect the voltage and active and reactive power dynamic of the WT.

2.1.1. WT Control Mode

In terms of the active power and reactive power output of the WT, the control methods of the WTs are mainly divided into constant voltage control and constant power factor control [19,20]. The constant voltage control mode of the WT is shown in the Q local control in Figure 1. Under this control mode, the active power fluctuation at the grid connection point is large and the fault traversal ability is low when the WT fails. The constant power factor control of the WT is shown in the V local control in Figure 1. Under this control mode, the WT can enter the steady state quickly after the power grid failure, but the active power recovery value is usually too large. The specific formula of WT is too complex and is not the focus of this paper, so it is not discussed here.

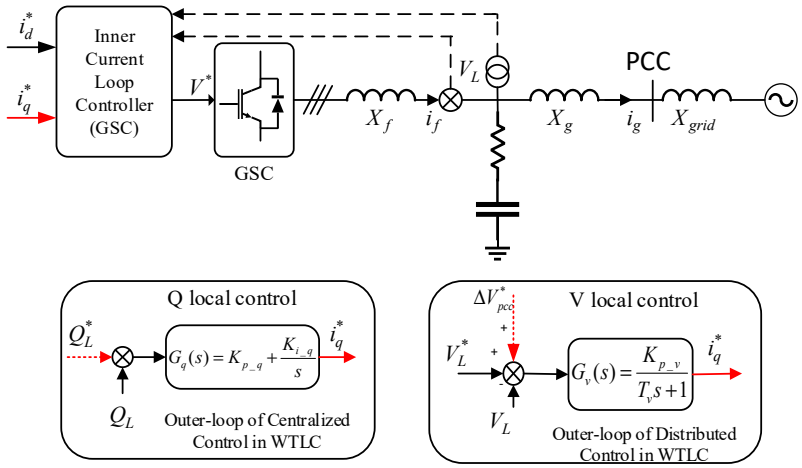


Figure 1. The control mode of the WT.

2.1.2. Wind Speed and Its Wake Effect

The WT generally adopts the maximum power tracking control method, so the wind speed is the most important factor affecting the active power output of the WT. For example, the wind speed in the offshore WF is more uniform than that in the onshore WF, the wake effect has a greater impact, and the huge economic loss caused by the wake effect every year [21,22]. At the same time, the downstream WTs in the WF are shielded by the upstream WTs, and their input wind speed is lower than that of the upstream WTs. The relationship can be described as the wake effect, as can be seen in Figure 2. Due to the existence of the wake effect, the dynamic characteristics of active power of different WTs are no longer independent of each other.

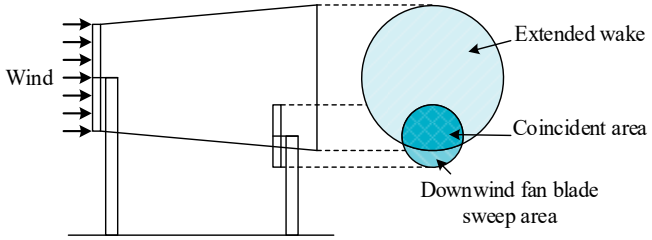


Figure 2. Scheme of overlapped wake model.

2.1.3. Influence of Wind Direction on Wake Effect

When the wind direction changes, the windward side of the WT blade is always perpendicular to the wind direction, as shown in Figure 3. As the wind direction changes, the wake effect between the WTs changes accordingly. The change of the wake effect between the WTs would cause the changes of the rotational speed and pitch angle of the WTs, and at the same time, it would change the overall active power output of the WF.

(**a**) (**b**)

Figure 3. (**a**) Wind direction is 0°. (**b**) Wind direction is 16.5°.

2.1.4. Resistance Value of Crowbar Protection Resistor

At present, doubly-fed induction generators (DFIG) are generally equipped with crowbar protection. When there are power grid faults, the connection of the crowbar protection would lock the rotor-side converter of the WT, causing the WT to run asynchronously to absorb a large amount of reactive power from the power grid; thereby, the dynamic characteristics of reactive power of the DFIG during low voltage ride-through changes significantly [23,24].

2.1.5. The Degree of Voltage Drop at the WT Port

According to the principle of short-circuit calculation in power system analysis, the calculation formula of short-circuit point current and WT port voltage after short-circuit is:

$$\begin{cases} \dot{I}_f = \frac{\dot{U}_{f|0|}}{Z_{ff}+z_f} \\ \dot{U}_i = \dot{U}_{i|0|} - Z_{if}\dot{I}_f \end{cases} \quad (1)$$

where $\dot{I}_f$ represents the short-circuit current of the short-circuit point f; $\dot{U}_i$ represents the voltage of the WT port i after the short-circuit; $\dot{U}_{f|0|}$ and $\dot{U}_{i|0|}$ represent the steady-state voltage of the short-circuit point f and the WT port i before the short-circuit, respectively; and Z_{ff}, Z_{if}, and z_f represent the self-impedance of the short-circuit point f, the mutual impedance between the WT port i and the short-circuit point f, and the short-circuit impedance, respectively.

It can be seen from (1) that the WT port voltage after a power system fault is mainly related to the short-circuit current at the short-circuit point and the mutual impedance between the WT port and the short-circuit point, and the short-circuit current at the short-circuit point is related to the short-circuit point self-impedance. Therefore, the change of fault location (directly related to the degree of voltage drop at the WT port) would cause the difference of dynamic voltage at the WT port, which would affect the dynamic characteristics of active and reactive power output of the WT.

2.2. Relationship between Key Factors and WF Equivalence

From the above analysis, it is known that there are many factors that affect the dynamic characteristics of the WT output. The purpose of this paper is to obtain the mathematical expression of the dynamic characteristics of the WT output through data mining, which is to obtain effective information that can describe the dynamic characteristics of the WT from a large number of dynamic characteristics of the WT output, and use this as a clustering indicator for WF equivalence. Among the key factors affecting the dynamic characteristics of the WT output, some factors are constant for a WT, such as the resistance of the crowbar protection resistor; but more factors are changeable, such as wind speed. The difference in the dynamic characteristics of the WT output is also mainly due to the changes of these random key factors. Therefore, by adjusting the value of the random key factor and taking it as the input of the WT, a large number of output dynamic characteristics of the WT can be obtained, which provides the necessary basis for the subsequent data mining work.

3. Extraction Method of WT Clustering Indicator Based on MMSE

MMSE is briefly reviewed, and the applicability analysis of MMSE in WF equivalence is presented. Further, the extraction process of WT clustering indicator is given.

3.1. MMSE Overview

MMSE is established on the basis of multivariate sample entropy proposed by Ahmed and is mainly used to evaluate and quantify multivariate time series complexity [25,26]. MMSE overcomes the influence of time series length on entropy calculation, and shows great advantages in operating any number of data channels at the same time. The acquisition of MMSE mainly consists of two steps: time series coarsening and multivariate sample entropy calculation.

3.1.1. Processing of Time Series Coarsening

For the p-dimensional time series $\{y_{kl}\}_{l=1}^{L}(k = 1, 2, \ldots, p)$, L is the length of the time series, and the time series after coarsening can be expressed as:

$$x_{k,i}^{\lambda} = \frac{1}{\lambda} \sum_{l=\lambda(i-1)+1}^{\lambda i} y_{k,l} \; (1 \le i \le [\frac{L}{\lambda}] \; , \; k = 1, \; 2, \; \ldots, \; p) \tag{2}$$

where the square brackets represent the round-down symbol; λ represents the adjustable coarse-grained scale factor; and $x_{k,i}^{\lambda}$ represents the element of the time series k after λ coarse graining. Scale factor λ is the most important parameter to adjust the coarsening degree of time series in MMSE. Generally speaking, λ is an integer greater than or equal to 1, for each λ, the complexity evaluation of time series under this scale factor can be obtained. By adjusting the value of λ, the evaluation of time series under multi-scale factor can be obtained.

On the one hand, time series coarsening reduces the influence of time series length on the calculation of entropy value by averaging. On the other hand, the complexity description of time series on multiple time scales is achieved by coarsening, which greatly reduces the influence of outliers on time series complexity calculation.

3.1.2. Calculation of MMSE

The calculation of multivariate sample entropy is the basis of MMSE. The calculation of multivariate sample entropy can be divided into five steps, which are described in detail below.

(1) Set the embedding vector $\boldsymbol{M} = [m_1, m_2, \ldots, m_p] \in \mathbf{R}^P$, the time delay vector $\boldsymbol{\tau} = [\tau_1, \tau_2, \ldots, \tau_p]$. Let $N = [L/\lambda]$, and the vector reconstruction of p-dimensional time series $\left\{x_{k,i}^{\lambda}\right\}_{i=1}^{N}$ after λ coarsening is performed to obtain m-dimensional composite delay vectors, and the number of the vectors is $(N-v_m)$:

$$\begin{aligned} X_m^{\lambda}(i) = [&x_{1,i}^{\lambda}, x_{1,i+\tau_1}^{\lambda}, \ldots, x_{1,i+(m_1-1)\tau_1}^{\lambda}, x_{2,i}^{\lambda}, x_{2,i+\tau_2}^{\lambda}, \ldots, x_{2,i+(m_2-1)\tau_2}^{\lambda}, \\ &x_{p,i}^{\lambda}, x_{p,i+\tau_p}^{\lambda}, \ldots, x_{p,i+(m_p-1)\tau_p}^{\lambda}] \ (i = 1, 2, \ \ldots, N - v_m) \end{aligned} \tag{3}$$

where $m = \sum_{k=1}^{p} m_k$ and v_m is defined as $v_m = \max\{\boldsymbol{M}\} \times \max\{\boldsymbol{\tau}\}$.

(2) For any two composite delay vectors $X_m^{\lambda}(i)$ and $X_m^{\lambda}(j)$, find the maximum norm between the two vectors, denoted as $d[X_m^{\lambda}(i), X_m^{\lambda}(j)]$. Set the threshold r, and count the number of vector pairs with $d[X_m^{\lambda}(i), X_m^{\lambda}(j)] \le r$ $(1 \le j \le N-v_m, j \ne i)$, denoted as P_i^m. Calculate the frequency $B_i^m(r)$ of the vector pairs that meet the above conditions, and calculate the average value to obtain $B_m(r)$. The calculation formulas of the two are as follows:

$$\begin{cases} B_i^m(r) = \frac{1}{N-v_m-1} P_i^m \\ B_m(r) = \frac{1}{N-v_m} \sum\limits_{i=1}^{N-v_m} B_i^m(r) \end{cases} \tag{4}$$

(3) Add the value 1 to any element m_k in the above embedding vector $\boldsymbol{M}$ to obtain a new embedding vector $\boldsymbol{M}'_k = [m_1, m_2, \ldots, m_k{+}1, \ldots, m_p]$. Replace $\boldsymbol{M}'_k$ with $\boldsymbol{M}$ in step (1), and use the method of step (1) to obtain a new $m+1$ dimensional composite delay vector $X_{m+1}^{\lambda,k}$. By changing k from 1 to p, $\sum_{k=1}^{p}(N-v_{m+1}^k)$ total of $X_{m+1}^{\lambda,k}$ can be obtained, where $v_{m+1}^k = \max\{\boldsymbol{M}'_k\} \times \max\{\boldsymbol{\tau}\}$.

(4) For any two $X_{m+1}^{\lambda,k}(i)$ and $X_{m+1}^{\lambda,k}(j)$, the calculation satisfies the condition $d[X_{m+1}^{\lambda,k}(i), X_{m+1}^{\lambda,k}(j)] \le r$ $(1 \le j \le \sum_{k=1}^{p}(N-v_{m+1}^k), j \ne i)$, denote the number of vector pairs as P_i^{m+1}, and calculate $B_i^{m+1}(r)$ and $B_{m+1}(r)$, the two calculation formulas are, respectively, as follows:

$$\begin{cases} B_i^{m+1}(r) = \frac{1}{\sum_{k=1}^{p}(N-v_{m+1}^k)-1} P_i^{m+1} \\ B_{m+1}(r) = \frac{1}{\sum_{k=1}^{p}(N-v_{m+1}^k)} \sum\limits_{i=1}^{\sum_{k=1}^{p}(N-v_{m+1}^k)} B_i^{m+1}(r) \end{cases} \tag{5}$$

(5) Calculate multivariate sample entropy after λ coarsening. By changing the value, the MMSE at different scales can be obtained.

$$E(\lambda, \boldsymbol{M}, \boldsymbol{\tau}, r, N) = -\ln\left[\frac{B_{m+1}(r)}{B_m(r)}\right] \tag{6}$$

By coarsening a time series with a scale factor and calculating the multivariate sample entropy of the processed time series, the multivariate sample entropy at that scale can be obtained. By selecting different scale coefficients, the multivariate sample entropy at multiple scale coefficients, which is MMSE, can be obtained.

3.2. Applicability Analysis of MMSE in WF Equivalence

MMSE is an effective tool to deal with the complexity of multivariate time series and is essentially a data mining method. In Section 2, the key factors affecting the dynamic characteristics of the WT output are mainly analyzed. However, the mathematical relationship between these factors and the dynamic characteristics of the WT output is difficult to obtain, which is to say, it is difficult to establish the relationship between the key factors and the dynamic characteristics of the WT output.

The output dynamic characteristics of the WT are affected by many factors and are not unique. Therefore, the dynamic characteristics of the WT are random and fuzzy. As an effective tool to characterize the degree of uncertainty of the system in science and technology, entropy can describe the randomness and fuzziness in the dynamic characteristics of WTs. In the simulation, the dynamic characteristics of the WT's active power output are usually expressed as time series. Yet, the classical entropy theory is difficult to deal with the time series. Therefore, the MMSE is introduced to describe the randomness and fuzziness of the time series. At the same time, a scale factor is introduced to obtain a multi-scale complete description of the dynamic characteristics of the power output by the WT.

3.3. Extraction Process of WT Clustering Indicator

Among the key factors affecting the dynamic characteristics of the WT output, the WT control method, wake effect under defined wind direction, and the resistance value of the crowbar protection resistor may be different in different WTs, but they are fixed for a single WT; the wind speed, wind direction, and grid short-circuit fault location are random and cannot be exhaustive. Therefore, it is necessary to count the probability characteristics of these random key factors based on historical data, sample them by constructing a random variable generator, and input them into the WT. Taking the sampling of wind speed as an example, if the wind speed at the location of the WF is a normal distribution through the historical statistics, a certain number of wind speed samples can be obtained by random sampling. The sampling of wind direction and fault location is exactly the same as that of wind speed.

Input the sampling samples into the WT, and the WT power output curve with the same number of sampling samples can be obtained. Select coarsening scale factor λ, and use formula (2) to calculate the coarsening results of the WT active power output curve under this scale factor $x_{k,}^{\lambda} I$ ($1 \leq i \leq [L/\lambda]$, $k = 1, 2, \ldots, p$). Through adjusting the size of λ, different coarsening results can be obtained. For the coarse-grained result of the WT's active power output curve under a certain scale factor λ, the multivariate sample entropy after λ coarse-grained can be obtained by using Equations (3)–(6), which is, $E(\lambda, \boldsymbol{M}, \boldsymbol{\tau}, r, N)$. In this paper, the multivariate sample entropy of WT active power output curve under different scale factor λ, namely MMSE, is taken as the clustering indicator of WT. The extraction process of WT clustering indicator can be shown in Figure 4.

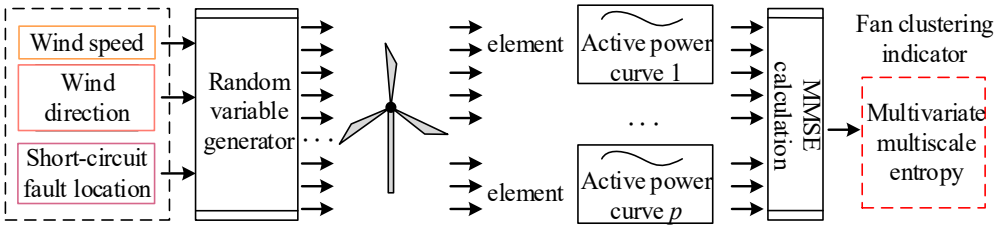


Figure 4. Extraction process of WT cluster indicator.

4. WF Equivalent

The WT clustering method based on MV-FCM is presented, and the calculation of the equivalent WT parameters is given.

4.1. WT Clustering Method Based on MV-FCM

In this paper, the MMSE of active power and reactive power is considered as the clustering indicator. Apparently, the MMSEs of active power and reactive power consist of two views [27], and traditional single-view clustering algorithms cannot be applied. In recent years, multi-view clustering has become a hot topic in data mining. These methods can aggregate data samples with multiple views into several groups, and each view has its own feature. In [28], the researchers proposed a MV-FCM clustering method, which firstly calculates the membership degree for each view, and then combines the clustering results of all views into the final clustering result.

Assuming that the dataset can be divided into C classes, the objective function and constraints of the MV-FCM clustering algorithm can be expressed as (7).

$$\min \sum_{k=1}^{K} \sum_{j=1}^{N} \sum_{i=1}^{C} \begin{matrix} \{\mu_{ij,k}^{m} d_{ij,k}^{2} + \alpha_{j,k} \mu_{ij,k} \left(1 - \mu_{ij,k}^{m-1}\right) \\ -\beta_{j,k} \mu_{ij,k} (1 - \mu_{ij,k}^{m-1})\} \end{matrix} \tag{7}$$

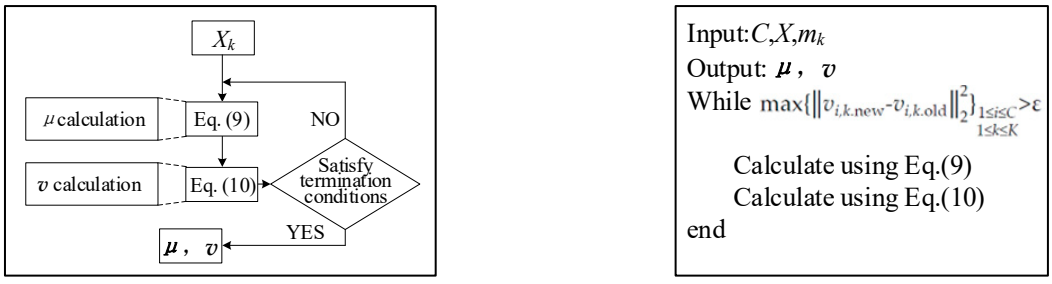
$$\begin{matrix} \text{s.t.} \sum_{i=1}^{C} \mu_{ij,k} = 1 \; \mu_{ij,k} \in [1,0] \\ 1 \le i \le C \; \le j \le N, \; 1 \; \le \; k \; \le \; K \end{matrix} \tag{8}$$

where N is the number of samples; K is the number of viewing angles; $\mu_{ij,k}$ is the membership degree of sample $\boldsymbol{x}_{j,k}$ to cluster center $v_{i,k}$ in perspective K; $d_{ij,k}^{2}$ is the Euclidean distance between $\boldsymbol{x}_{j,k}$ and $v_{i,k}$; the role of $\alpha_{j,k}\mu_{ij,k}\left(1 - \mu_{ij,k}^{m-1}\right)$ is to make the membership degree of $\boldsymbol{x}_{j,k}$ to $v_{i,k}$ more important in clustering, while $\beta_{j,k}\mu_{ij,k}\left(1 - \mu_{ij,k}^{m-1}\right)$ is to weaken this role. $\alpha_{j,k}$ and $\beta_{j,k}$. are enhancement parameters and weakening parameters, respectively; $m > 1$ is the fuzzification constant. The iteration update formula for $\mu_{ij,k}$ and $v_{i,k}$ is

$$\mu_{ij,k} = \frac{1}{\sum_{i=1}^{C} \left[\frac{d_{ij,k}^{2} - (\alpha_{j,k} - \beta_{j,k})}{d_{ij,k}^{2} - (\alpha_{j,k} - \beta_{j,k})} \right]^{\frac{1}{m-1}}} \tag{9}$$

$$v_{i,k} = \frac{\sum_{j=1}^{N} \mu_{ij,k}^{m} \boldsymbol{x}_{j,k}}{\sum_{j=1}^{N} \mu_{ij,k}^{m}} \tag{10}$$

The clustering process of MV-FCM is shown in Figure 5.



(**a**) (**b**)

Figure 5. Clustering process of MV-FCM. (**a**) Flow chart, and (**b**) pseudocode.

The quantification formula for the clustering effect is as follows:

$$\begin{matrix} CE = \frac{\left\{\sum_{i=1}^{k} [(\sum_{n=1}^{k_i} |E_{i,n} - E_{i,n}|)/k_i]\right\}/k}{(\sum_{i=1}^{k} |C_i - C_i|)/k} \\ \left(n = 1, 2, \ldots, k_i, \; n \; \neq \; n \; ; i = 1, 2, \ldots, k, i \neq i\right) \end{matrix} \tag{11}$$

where *CE* indicates the clustering effect, and the smaller the value, the better the clustering effect. $E_{i,n}$ represents the n-th clustering sample of class i. This paper refers to the multi-scale and multi-scale entropy of the WT. C_i represents the cluster center of category i; k_i represents the number of samples for category i; and k represents the number of categories.

4.2. Calculation of Equivalent WT Parameters

Assuming that the model and capacity of WTs in the WF are the same, the equivalent WT parameters are related to the number of WTs m. After the WT clustering is completed, the relevant parameters of the equivalent WT can be calculated by Equation (12):

$$\begin{cases} S_{\text{eq}} = \sum\limits_{i=1}^{m} S_i, P_{\text{eq}} = \sum\limits_{i=1}^{m} P_i, Q_{\text{eq}} = \sum\limits_{i=1}^{m} Q_i \\ x_{m\text{-eq}} = \frac{x_m}{m}, x_{s\text{-eq}} = \frac{x_s}{m}, x_{r\text{-eq}} = \frac{x_r}{m} \\ r_{s\text{-eq}} = \frac{r_s}{m}, r_{r\text{-eq}} = \frac{r_r}{m} \\ H_{\text{eq}} = \sum\limits_{i=1}^{m} H_i, K_{\text{eq}} = \sum\limits_{i=1}^{m} K_{\text{i}}, D_{\text{eq}} = \sum\limits_{i=1}^{m} D_{\text{i}} \end{cases} \tag{12}$$

where the subscript eq represents the equivalent parameter; m refers to the number of equivalent units; S, P and Q, respectively, represent the capacity, active power and reactive power of the WT; x_m refers to reactance of excitation branch; x_s and r_s, respectively, represent reactance and resistance of stator winding; x_r and r_r, respectively, represent reactance and resistance of rotor winding; H, K and D, respectively, represent the shafting inertia time constant, shafting stiffness coefficient and shafting damping coefficient.

In this paper, the WT is connected to the public connection point through the step-up transformer. Since the WT capacity is the same, it can be assumed that the step-up transformer capacity is also the same, and the equivalent parameters of the transformer can be calculated by (13).

$$\begin{cases} S_{\text{T-eq}} = mS_{\text{T}} \\ x_{\text{T-eq}} = \frac{x_T}{m} \end{cases} \tag{13}$$

where S_{T} represents transformer capacity; x_T represents transformer reactance. Based on the principle of constant voltage loss before and after equalization, the line impedance is equalized, and the equivalent formula is shown in (14):

$$\begin{cases} Z_{\text{eq}} = \frac{\sum_{i=1}^{m}\left(\sum_{k=1}^{i}(Z_k \sum_{j=k}^{n} P_j)P_i\right)}{\left(\sum_{i=1}^{m} P_i\right)^2} \\ Y_{\text{eq}} = \sum\limits_{\text{i}=1}^{m} Y_i \end{cases} \tag{14}$$

where n represents the number of WTs in the main-line WT branch; Z_k represents the branch impedance of the k-th main-line cable; P_i represents the output power of the i-th WT; and Y_i represents the ground admittance of the i-th main-line cable.

5. Simulation Examples

5.1. Example Description

In this paper, PSCAD simulation software is used to build the IEEE14 node model including the WF shown in Figure 6. The WF consists of 16 WTs, numbered W1–W16, with a single WT capacity of 1.5 MW. The WF is connected to a common connection point PCC through a box transformer (660 V/35 kV) and a collector line, and then to the power system through a main transformer (35 kV/110 kV). The parameters of the WT are shown in Table 1. It is assumed that the wind speed obeys a Weibull distribution with a scale coefficient of 10.7 and a shape coefficient of 3.97 [29], the WT control mode obeys a two-point distribution, and the location of faults in each line obeys a uniform distribution. For the wind direction, 0°–360° is divided into 16 wind direction zones, and the span of each wind direction zone is 22.5°. The calculation results considering the wake effect of the wind direction in Figure 5 are shown in Table 2. The numbers in the table represent the ratio of the wind speed of the

input WT to the natural wind speed. The crowbar protection resistance of W1–W8 is 0.14 Ω, and the crowbar protection resistance of W9–W16 is 0.12 Ω. The system configuration used in the simulation is Intel (R) Core (TM) i7—7700 CPU 3.60 GHz, 16 GB memory.

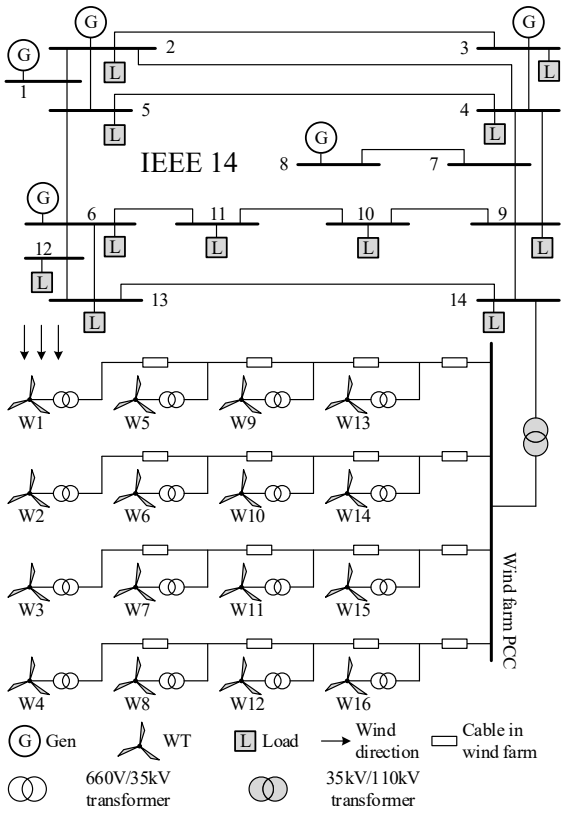


Figure 6. Simulation system.

Table 1. Parameters of WT.

WT	Gen	Converter
Rated capacity = 1.5 MW cut-in wind speed = 4.5 m/s cut-out wind speed = 18 m/s	R_s = 0.00706 p.u. R_r = 0.005 p.u. L_m = 2.9 p.u. L_s = 0.171 p.u. L_r = 0.156 p.u.	Grid side inductance = 0.15 p.u. Grid side resistance = 0.0015 p.u. DC capacitor = 0.01 F

Table 2. Calculation results of wake effect under a wind direction.

WT Number	Wake Effect	WT Number	Wake Effect	WT Number	Wake Effect	WT Number	Wake Effect
1	1.000	5	1.000	9	1.000	13	1.000
2	0.947	6	0.945	10	0.944	14	0.952
3	0.908	7	0.924	11	0.915	15	0.906
4	0.829	8	0.878	12	0.856	16	0.904

5.2. Equivalent Model of WF

Consider the random combination of wind speed between 5 m/s and 18 m/s, wind direction between 0°, 90°, 180°, 270°, and short-circuit fault location between nodes 1–14 to form different simulation conditions. The simulation system shown in Figure 6 is simulated and analyzed. Among them, it is assumed that the fault occurs at t = 3 s, and the fault duration is 0.15 s. Taking the active power curve of the WT outlet as the analysis object, Figure 7 shows the active power curve at the outlet of W1 under various random combinations, which is the data source for extracting the multivariate multi-scale entropy of

the active and reactive power of W1. In Figure 7, each curve is an "element" corresponding to the multivariate multi-scale entropy. Calculate the multivariate multi-scale entropy of all the curves in Figure 7 to obtain the multivariate multi-scale entropy of active power and reactive power, that is, the multi-view clustering indicator of W1. The maximum scale factor is selected to be 10, and the multi-scale entropy value of the W1–W16 active and reactive power curves is calculated using the WT clustering indicator extraction method in Section 3 of this paper. The results are shown in Figure 8. The vector in the two figures is the multi-view clustering indicator of each WT. The difference in entropy value of different WTs at multiple scales reflects the difference in the dynamic characteristics of the WT output. Taking the multi-scale entropy value of the active and reactive power of the WT as the multi-view clustering indicator, the MV-FCM algorithm is used for clustering, where the number of clusters is four, and the clustering results are shown in Table 3.

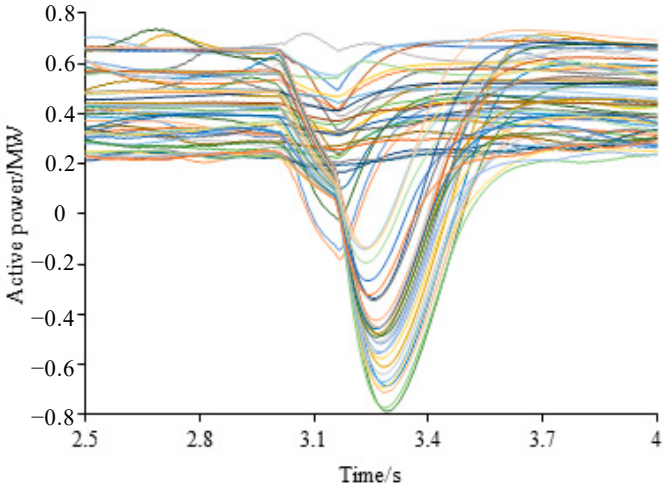
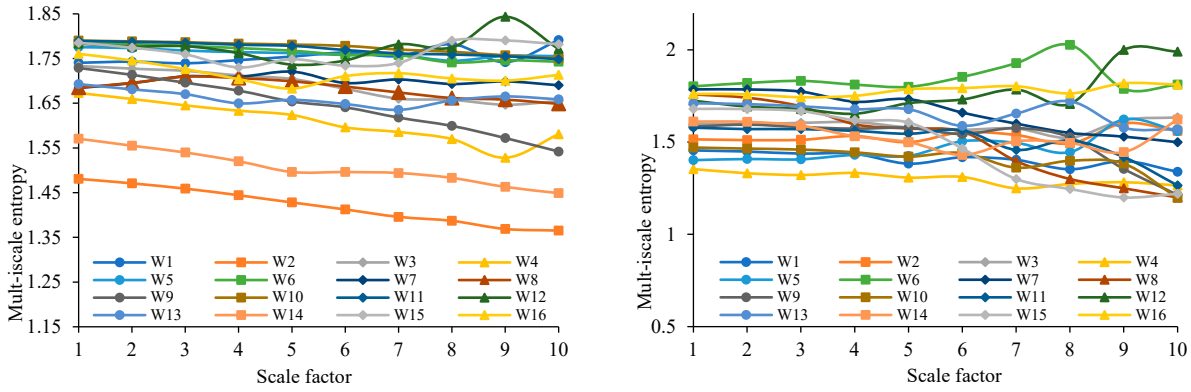


Figure 7. Active power of W1 under different random combination.



(**a**) (**b**)

Figure 8. (**a**) Multi-scale entropy of active power in W1–W16. (**b**) Multi-scale entropy of reactive power in W1–W16.

Table 3. Cluster of wind generators.

Cluster	Wind Generators
Equivalent Cluster 1	W1, W5, W6, W10, W1, W15
Equivalent Cluster 2	W4, W91, W12
Equivalent Cluster 3	W3, W8, W13, W16
Equivalent Cluster 4	W2, W14, W7

5.3. Comparative Analysis of Equivalent Model and Detailed Model

In order to verify the accuracy of the equivalent model of the WF and the accuracy of the output dynamic characteristics of the equivalent WF in various scenarios, such as active power and reactive power, different fault situations of the WF are set, including different wind speeds, wind directions and fault locations, as shown in Table 4. It is also set that the fault occurs at t = 3 s, and the fault duration is 0.15 s. Under the operating conditions of Table 4, the active power, reactive power and voltage curves of the equivalent model and the detailed model at the PCC are shown in Figures 9 and 10.

Table 4. Operating conditions of WF.

Case	Wind Speed(m/s)	Wind Direction	Fault Location
1	9.5	North	Bus9
2	10	South	the middle of Bus4–Bus7
3	10.5	Southeast	the middle of Bus10–Bus11
4	11	East	the middle of Bus4–Bus5
5	11.5	Northwest	the middle of Bus2–Bus3
6	12	Northeast	the middle of Bus13–Bus14

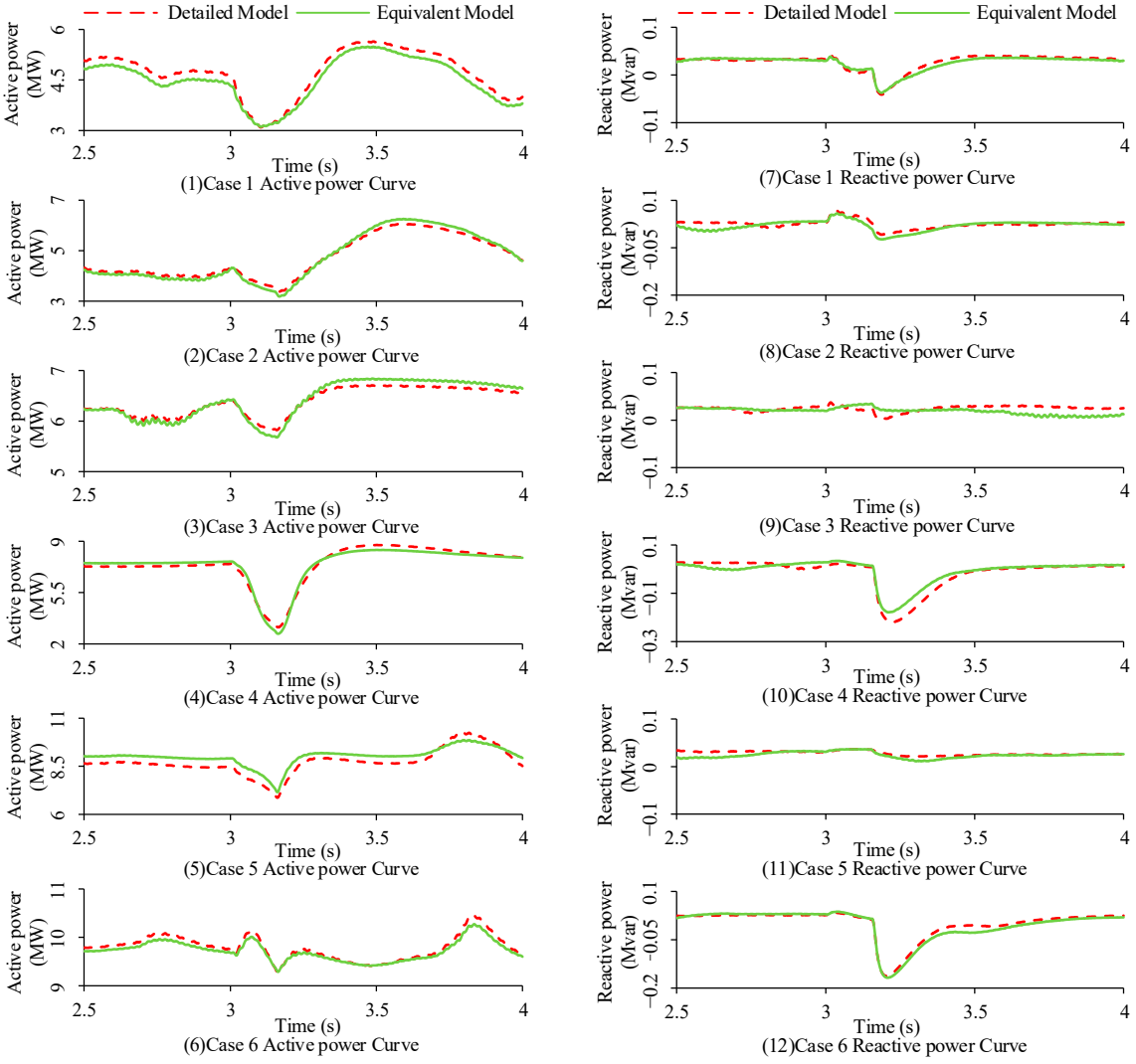


Figure 9. (**a**) Active power curves in PCC of different models. (**b**) Reactive power curves in PCC of different models.

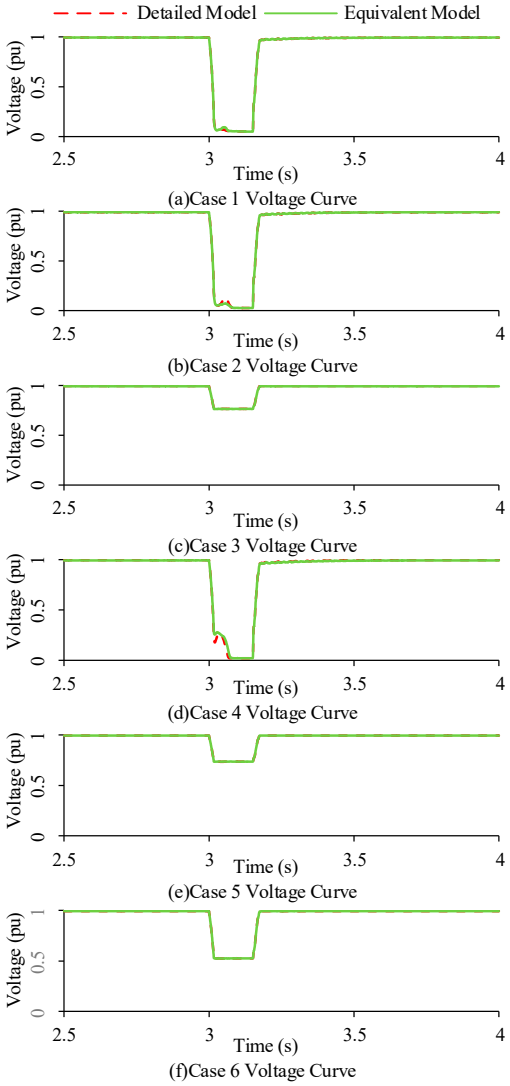


Figure 10. Voltage curves in PCC of different models.

As can be seen from Figures 9 and 10, under the conditions of WF faults in multiple scenarios such as different wind speeds, wind directions and fault locations, the active power, reactive power and voltage output characteristics of the dynamic equivalent model in this paper trace the detailed model well. Especially after the fault of the power system, the dynamic equivalent model of this paper can well reflect the dynamic characteristics of the WT, which verifies the accuracy of the output dynamic characteristics of the WF equivalent model in multiple scenarios. This is because the output dynamic characteristics of WTs under the influence of various key factors are comprehensively considered when extracting the clustering indicator of WTs in the WF. It can reflect the randomness and fuzziness of the WT under multi-scenario fault conditions, rather than just describe the output dynamic characteristics of the WT under a certain fault condition. At the same time, the multi-view clustering indicator is used for clustering, and the clustering results can reflect the dynamic characteristics of physical quantities from each viewpoint. Therefore, the clustering results are suitable for the simulation analysis of external faults of WF in various scenarios.

To further verify the validity of this model, set different number of clusters, and calculate the absolute average error E_{rse} and root mean square error E_{mres} of the voltage of

each node of the power system during the simulation period. The calculation formulas of both E_{rse} and E_{mres} are as follows:

$$\begin{cases} E_{rse} = \frac{1}{K}\sum_{j=1}^{K}\left|V_{i,j} - V'_{i,j}\right| \\ E_{mres} = \sqrt{\frac{\sum_{j=1}^{K}\left(V_{i,j} - V'_{i,j}\right)^2}{K}} \end{cases} (j = 1, 2, \ldots, G) \quad (15)$$

where K represents the number of sample points of the simulation curve; G represents the number of nodes in the power system; $V_{i,j}$ and $V'_{i,j}$ represent the results of equivalent model and detailed model, respectively.

Figure 11 shows E_{rse} and E_{mres} of IEEE14 system node voltage under WF equivalent model and detailed model with different number of clusters. From the figure, it can be seen that the model error does not continuously decrease with the increase of the number of equivalent machines, among which the simulation error of four machine equivalents is the smallest. This is because in the clustering process, the optimal clustering effect is to ensure that the distance between the same class of samples is as small as possible, and the distance between different cluster centers is as large as possible.

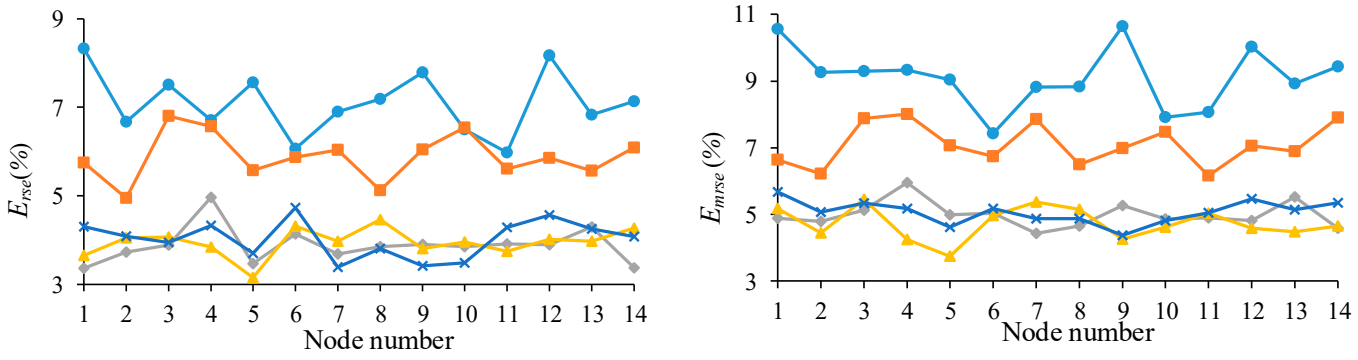


Figure 11. (**a**) E_{rse}% of different equivalent WT numbers (clustering); (**b**) E_{mres}% of different equivalent WT numbers (clustering). Where light blue with circular line represents 2 machine equivalent, orange with square line represents 3 machine equivalent, gray with diamond line represents 4 machine equivalent, yellow with triangle line represents 5 machine equivalent, dark blue with × line represents 6 machine equivalent.

Table 5 shows the *CE* values for different cluster numbers. It can be seen from the table that when the number of clusters is 4, the *CE* value is the smallest, so the simulation error of the 4-machine equivalent is the smallest. When the number of clusters is less than 4, the number of clusters is too small. The result is that distance between clustering objects in the same category is too large, and the molecular term in Formula (12) is too large, which results in a larger *CE* value and poor clustering effect of WT. When the number of clusters is more than four, small distances between different cluster centers occur, and the denominator term in Formula (12) is too small, which also results in larger *CE* and poorer clustering effect of WT.

Table 5. E_{rse} and E_{mres} of voltage in PCC.

Cluster Number	2	3	4	5	6
CE	0.86	0.57	0.14	0.28	0.32

5.4. Comparison and Analysis between the Model in This Paper and Traditional Model

In this paper, multivariate multi-scale entropy is used for the first time to mine the dynamic characteristics of WT active power time series output. However, previous studies

mostly used single time point physical quantities such as wind speed and WT internal state quantity as clustering indicators, and the comparison between time series and single time points is to be studied. However, in the traditional method, the clustering indicator is usually extracted from a certain operation of the WT, not from the various fault conditions proposed in this paper. Based on this idea, as a comparison example in this section, this paper selects the active power time series of a wind in a fault situation (wind speed 10 m/s, wind direction south wind, short circuit fault location Bus 9 of power grid) as the object to be excavated, and extracts its MMSE as the clustering indicator. It should be noted that since only a single case is considered, MMSE would degenerate into a unit multi-scale entropy. Table 6 shows E_{rse} and E_{mres} of the voltage at the outlet of the WF under different fault conditions when two clustering indices are selected. From the table, it can be seen that the comparison model only has small error when extracting the corresponding fault case as clustering indicator, and the errors in non-clustering cases are larger than the model in this paper. This is because when extracting clustering indicator, the model takes into account the dynamic characteristics of the WT output affected by various key factors comprehensively, so it has a wide applicability. The comparison model only considers a single fault when choosing the clustering indicator, so its simulation under this fault condition has better adaptability, while there are larger simulation errors in other cases.

Table 6. E_{rse} and E_{mres} of voltage in PCC.

Fault Condition	E_{rse} (%)		E_{mres} (%)	
	Proposed Model	Comparison Model	Proposed Model	Comparison Model
10 m/s, South wind, Bus 9	4.33	3.77	4.45	4.15
10 m/s, North wind, Bus 9	4.26	8.29	1.65	8.66
11 m/s, North wind, Bus 9	4.29	9.06	4.67	9.78
10 m/s, South wind, Bus 4	4.42	7.84	4.54	7.29

5.5. Simulation Efficiency Analysis of Equivalent Model

In order to verify the superiority of the equivalent model in this paper in terms of simulation efficiency, several fault situations are set up, and the simulation time of the equivalent model and the detailed model is obtained through simulation, as shown in Table 7. From this table, compared to the detailed model, the simulation time of the equivalent model in this paper is reduced by about 86%.

Table 7. Simulation time comparison of equivalent model and detail model.

Fault Condition	Simulation Time (s)	
	Proposed Model	Detailed Model
10 m/s, North wind, Bus 9	133	946
11 m/s, North wind, Bus 9	136	951
10 m/s, South wind, Bus 4	134	953

6. Conclusions

In this paper, an equivalent model of WF based on MMSE and multi-view clustering is proposed. Compared with the traditional WF equivalent model, this model can better simulate the dynamic process of WF in all kinds of fault scenarios of power systems, reducing the number of WF equivalence. Compared with detailed models, the proposed equivalent model can reduce the simulation time by about 86%. Compared with the comparison model, the proposed equivalent model improves the accuracy by about 44%, which greatly improves the simulation efficiency. The number of clusters is the factor that affects the equivalent value of WF. In the simulation case of this paper, the equivalent precision is higher when the number of equivalent machines is 4. The method proposed in this paper is only suitable for the equivalent modeling of a single WF. When modeling multiple WFs, the correlation of the active output of each WF becomes weak due to the different installation locations of some WFs affected by meteorological and geographic

interference factors. How to consider the MMSE method to establish the equivalent model of multiple WFs is the future research direction of this paper.

Author Contributions: Conceptualization, J.H. and L.L.; methodology, J.H.; software, H.S.; validation, J.H., L.L. and Y.Q.; formal analysis, Y.Q. and M.L.; investigation, H.S.; writing—original draft preparation, J.H. and Z.S.; writing—review and editing, J.H. and L.L.; visualization, J.H.; supervision, H.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by [National Natural Science Foundation of China] grant number [61773137] and [Natural Science Foundation of Shandong Province] grant number [ZR2019MF030].

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Council GWE. *GWEC Global Wind Report 2022*; Global Wind Energy Council: Bonn, Germany, 2022; pp. 119–125.
2. Ding, M.; Zhu, Q.; Han, P. Analysis on equivalent model for wind farms. *Smart Grid.* **2014**, *2*, 1–6.
3. Mahela, O.P.; Shaik, A.G. Comprehensive overview of grid interfaced wind energy generation systems. *Renew. Sustain. Energy Rev.* **2016**, *57*, 260–281.
4. Zou, J.; Chao, P.; Yan, Y.; Hong, Z.; Yan, L. A survey of dynamic equivalent modeling for wind farm. *Renew. Sustain. Energy Rev.* **2014**, *40*, 956–963.
5. Fernandez, L.M.; Garcia, C.A.; Saenz, J.R.; Jurado, F. Equivalent models of wind farms by using aggregated wind turbines and equivalent winds. *Energy Convers. Manag.* **2009**, *50*, 691–704.
6. Li, H.; Wan, Q.; Xiang, C. Wind farm equivalence method considering wind speed. *Electr. Power Autom. Equip.* **2013**, *33*, 121–123.
7. Xu, Y.; Liu, D. Equivalence of wind farms with DFIG based on two-step clustering method. *Power Syst. Prot. Control.* **2017**, *45*, 108–114.
8. Gao, Y.; Jin, Y.; Ju, P.; Zhou, Q. Dynamic Equivalence of Wind Farm Composed of Double Fed Induction Generators Considering Operation Characteristic of Crowbar. *Power Syst. Technol.* **2015**, *39*, 628–633.
9. Chen, S.; Wang, C.; Shen, H.; Gao, N.; Lan, H. Dynamic Equivalence for Wind Farms Based on Clustering Algorithm. *Proc. CSEE* **2012**, *32*, 11–19.
10. Gupta, A.P.; Mitra, A.; Mohapatra, A.; Singh, S.N. A Multi-Machine Equivalent Model of a Wind Farm Considering LVRT Characteristic and Wake Effect. *IEEE Trans. Sustain. Energy* **2022**, *13*, 1396–1407.
11. Zhang, X.; Li, L.; Hu, X.; Wang, X.; Zhou, X. Wind Farm Dynamic Equivalence Based on Clustering by Output Time Series Data of Wind Turbine Generators. *Power Syst. Technol.* **2015**, *39*, 2787–2793.
12. Zhou, M.; Ge, J.; Li, G. Cloud Model based DFIG Wind Farm Dynamic Voltage Equivalence Method. *Proc. CSEE* **2015**, *35*, 1097–1105.
13. Li, H.; Wang, H.; Shi, X.; Yang, C. Study on equivalent model of wind farms based on genetic algorithm. *Power Syst. Prot. Control.* **2011**, *39*, 1–8.
14. Wu, Q.; Zhong, Q.; Wang, G.; Li, H. Multi-machine Equivalent Approach of Wind Farm Based on Fuzzy C-means Clustering. *Mod. Electr. Power* **2016**, *33*, 27–32.
15. Zou, J.; Chao, P.; Xu, H.; Yan, Y. A Fuzzy Clustering Algorithm-Based Dynamic Equivalent Modeling Method for Wind Farm with DFIG. *IEEE Trans. Energy Convers.* **2015**, *30*, 1329–1337. [CrossRef]
16. Ali, M.; Ilie, I.S.; Milanovic, J.V.; Chicco, G. Wind Farm Model Aggregation Using Probabilistic Clustering. *IEEE Trans. Power Syst.* **2013**, *28*, 309–316. [CrossRef]
17. Lin, J. A unified scheme of grid operation index control based on big data platform. *Power Syst. Prot. Control.* **2018**, *46*, 165–170.
18. Bao, T.; Cheng, L.; Chen, B.; YU, T. Software design of reactive power optimization analysis based on artificial intelligence algorithms. *Power Syst. Prot. Control.* **2018**, *46*, 89–96.
19. Zhao, D.; Xu, R.; Zheng, L. Research on coordinated control strategy for DFIGs participating in system frequency regulation with different wind. *Power Syst. Prot. Control.* **2017**, *45*, 53–59.
20. Asadollah, S.; Zhu, R.; Liserre, M. Analysis of Voltage Control Strategies for Wind Farms. *IEEE Trans. Sustain. Energy* **2020**, *11*, 1002–1012. [CrossRef]
21. Barthelmie, R.J.; Jensen, L.E. Evaluation of wind farm efficiency and wind turbine wakes at the nysted offshore wind farm. *Wind. Energy* **2010**, *13*, 573–586. [CrossRef]
22. Nilsson, K.; Ivanell, S.; Hansen, K.S.; Mikkelsen, R.; Dan, H. Large-eddy simulations of the Lillgrund wind farm. *Wind Energy* **2014**, *18*, 449–467. [CrossRef]

23. Zou, Z.; Lei, Y.; Ran, O.; Zong, Z. Crowbar protection scheme based on dynamic resistance for reactive power improvement of doubly fed induction generator. *Power Syst. Prot. Control* **2018**, *46*, 87–94.
24. Qin, K.; Wang, S.; Kang, Z. Research on Zero-Voltage Ride through Control Strategy of Doubly Fed Wind Turbine. *Energies* **2021**, *14*, 2287. [CrossRef]
25. Ahmed, M.U.; Mandic, D.P. Multivariate multiscale entropy: A tool for complexity analysis of multichannel data. *Phys. Rev. E Stat. Nonlin. Soft Matter. Phys.* **2011**, *84*, 61918. [CrossRef] [PubMed]
26. Ahmed, M.U.; Mandic, D.P. Multivariate Multiscale Entropy Analysis. *IEEE Signal Processing Lett.* **2012**, *19*, 91–94. [CrossRef]
27. Jiang, Y.; Chung, F.; Wang, S.; Deng, Z.; Qian, P. Collaborative Fuzzy Clustering from Multiple Weighted Views. *IEEE Trans. Cybern.* **2015**, *45*, 688–701. [CrossRef]
28. Bezdek, J.C.; Ehrlich, R.; Full, W. FCM—The Fuzzy C-Means clustering-algorithm. *Comput. Geosci.* **1984**, *10*, 191–203. [CrossRef]
29. Han, J.; Miao, S.; Li, C. Probabilistic Optimal Energy Flow of Electricity-Gas-Heat Integrated Energy System Considering Correlation. *Trans. China Electrotech. Soc.* **2019**, *34*, 1055–1067.