sensors

MDPI

Article

Research on the Derated Power Data Identification Method of a Wind Turbine Based on a Multi-Gaussian–Discrete Joint Probability Model

Yuanchi Ma [1], Yongqian Liu [1,*], Zhiling Yang [1], Jie Yan [1], Tao Tao [1] and David Infield [2]

[1] State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, North China Electric Power University, Beijing 102206, China
[2] Wind Energy and Control Centre, Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow G1 1XQ, UK
* Correspondence: yqliu@ncepu.edu.cn

Abstract: This paper focuses on how to identify normal, derated power and abnormal data in operation data, which is key to intelligent operation and maintenance applications such as wind turbine condition diagnosis and performance evaluation. Existing identification methods can distinguish normal data from the original data, but usually remove power curtailment data as outliers. A multi-Gaussian–discrete probability distribution model was used to characterize the joint probability distribution of wind speed and power from wind turbine SCADA data, taking the derated power of the wind turbine as a hidden random variable. The maximum expectation algorithm (EM), an iterative algorithm derived from model parameters estimation, was applied to achieve the maximum likelihood estimation of the proposed probability model. According to the posterior probability of the wind-power scatter points, the normal, derated power and abnormal data in the wind turbine SCADA data were identified. The validity of the proposed method was verified by three wind turbine operational data sets with different distribution characteristics. The results are that the proposed method has a degree of universality with regard to derated power operational data with different distribution characteristics, and in particular, it is able to identify the operating data with clustered distribution effectively.

Keywords: wind turbine; SCADA data; derated power operation; multi-Gaussian–discrete joint probability model; EM algorithm

Citation: Ma, Y.; Liu, Y.; Yang, Z.; Yan, J.; Tao, T.; Infield, D. Research on the Derated Power Data Identification Method of a Wind Turbine Based on a Multi-Gaussian–Discrete Joint Probability Model. *Sensors* **2022**, *22*, 8891. https://doi.org/10.3390/s22228891

Academic Editor: Hossam A. Gabbar

Received: 13 September 2022
Accepted: 14 November 2022
Published: 17 November 2022

1. Introduction

Wind power has achieved large-scale development and utilization on a global scale and has become the most widely used and fastest-growing renewable energy source [1]. As an important part of the wind turbine, the Supervisory Control And Data Acquisition (SCADA) system provides detailed data on the operating status of the wind turbine [2–4]. In recent years, the wind power industry has developed rapidly, and wind farms have accumulated a large quantity of operational data. Such data are indispensable for wind turbine state assessment and wind power prediction. In real time, they are also important for power system dispatch and to schedule any wind farm curtailment or derating.

Generally, wind turbine power curves exhibit a degree of scatter, reflecting measurement uncertainty in both wind speed and power. In addition, a certain amount of abnormal data usually exists in the actual measured operational data of the wind farm, which complicates interpretation such as determination of wind turbine operating state or wind power prediction. There are many factors affecting the quality of operational data, such as the measurement error of the sensor itself, poor measurement accuracy caused by a poor operating environment, data storage and transmission error, wind turbine performance failure, and importantly, the operating of a wind turbine in a derated power state. It is useful to

divide abnormal operational data into two categories: often extensive data generated by the turbine under derated control; and a generally smaller amount of outlier data that deviate from the main data distribution, due to averaging over state transitions or some other sporadic factors [5–7]. However, there is no current SCADA parameter indicating whether it is derated or for the max power limit for a wind turbine. Therefore, how to identify a derated operation from within the turbine operational data and how to eliminate outliers are important research activities in the field of wind power.

Outlier detection is widely used in the field of wind power, and related research institutions have carried out significant research with useful results. There are many statistical methods for detecting outliers, which can be roughly divided into the following five categories: distribution-based outlier detection, depth-based outlier detection, cluster-based outlier detection, distance-based outlier detection, and outlier detection of density methods [8–12].

Distribution-based statistical outlier detection makes use of a fitted probability distribution for a given data set and identifies data that are far from this as outliers. The distribution-based outlier detection method is widely used in the wind power field. In [13], a mathematical model based on the quartile algorithm was used to identify the anomalous data. For the cases of a small amount of missing data, or in contrast, continuous missing blocks of data, the wind farm output correlation and multi-point cubic spline difference are used respectively for interpolation. The method reconstructs the missing data. In [14], the time series characteristics of bad data were identified, and a segmentation judgment method was applied. Any abnormal data are reconstructed based on the relationship between wind power output and the data characteristics of the wind farm itself. In [15], a joint probability model method based on the Copula function was proposed. By using the Copula function, a complex nonlinear multivariate relationship between parameters can be obtained based on the univariate marginal distribution of the data set. The significant outliers are then eliminated by examination of the derived joint probability model. In [16], an optimal intra-group variance algorithm for power curve analysis was proposed. This algorithm changes the dependence of traditional analysis methods on multi-dimensional data. It only needs to analyze wind speed and power and can identify the normal power generation status of the turbine. In [17], based on the analysis of the wind turbine-power abnormal operation data characteristics of wind turbines, the anomalous data are divided into four types: the bottom of the curve, the middle and upper stacking anomaly data, and the dispersive anomaly data around the curve. An anomalous data identification and cleaning process based on the combination of the change point grouping method and quartile method were proposed. This method can effectively identify four types of abnormal data, and the process is reasonable and the cleaning effect is good. The above distributed outlier detection methods can quickly and efficiently find outliers in the case of a known data set distribution. However, this method relies on the global distribution of a given data set and does not apply to situations where the high-dimensional data set and data set distribution are unknown.

In practice, most the operational data do not fully conform to a specific data model distribution [18,19]. To improve the distribution-based outlier detection method, the depth-based outlier detection method was created. This method assigns each data object a depth value and maps data objects to corresponding layers of a 2D space by the assigned depth values. Data objects on a shallow layer are more likely to be outliers than those on a deep layer. However, in practical applications, the existing depth-based outlier detection method is only effective in processing data in two-dimensional and three-dimensional space [20].

Cluster-based outlier detection divides the data set into clusters according to data features and identifies data points that are far away from any cluster as outliers. In [21], based on the k-means clustering, data stream concept drifting and existing outlier detection algorithm, a dynamic outlier detection algorithm was proposed. During the running of the algorithm, the sliding window size is adjusted adaptively according to the data flow concept drift detection result, and the cluster structure in the data set can be effectively

found while determining the outliers. Cluster-based outlier detection can identify outliers in a fast and timely manner. The detection of outliers is very sensitive to the clustering algorithm used; consequently, the clustering algorithm must be selected with care.

In order to improve the above-mentioned outlier detection method, researchers have proposed a distance-based outlier detection method. This approach assesses whether the distance of most data points in the data set from the target point is greater than the user-defined distance threshold, and if so, the target point is considered to be an outlier [22,23]. In [24], a fast distance-based data outlier detection algorithm was proposed. The algorithm uses the sliding window model to process the data stream and uses the vector inner product inequality to reduce the branch. The distance-based outlier detection method is widely used because of its simplicity and efficiency. However, since the method uses the global threshold and does not consider the local density change, only global outliers can be detected, and local outliers cannot be detected.

The density-based outlier detection method is built on the distance-based outlier detection method, and it determines an outlier based on the field of the data point. It is able to accurately find outliers with uneven data distribution. In [25], in order to improve the efficiency of the existing density-based outlier detection algorithms, an outlier detection algorithm based on local density, LDBO, was introduced. The concept of strong k nearest neighbor and weak k near point was introduced. By analyzing the outlier correlation of adjacent data points, individual data points are treated differently. A data point outlier pre-judgment strategy was proposed to effectively improve the efficiency of the outlier detection algorithm for data distribution anomalies. The density-based outlier detection method can solve the problem of local outlier detection well, but there are still problems of high complexity and parameter selection.

The above-mentioned outlier detection methods each have advantages and disadvantages, and their scope of application is different. To overcome the problems caused by a single method, the current research mostly adopts a mixture of two or more methods. There are outlier detection methods based on distribution and clustering. In [26], based on the analysis of the data characteristics of the identified wind outliers, the outlier data combination detection model based on quartile method and k-means clustering analysis was proposed. The model does not rely on the normal data set for training and learning. It has strong automated processing capability and versatility, but the k value of the method is more complex and has a greater impact on the data processing results. There are outlier detection methods based on distribution, clustering and density mixing. In [27,28], two quartile algorithms were used to eliminate sparse outliers, and then the DBSCAN (Density-Based Spatial Clustering of Applications with Noise) algorithm was used to eliminate stacked outliers. The method does not need to input the number of clusters, and it has high accuracy and universality. However, when the spatial clustering density is not uniform and the cluster spacing difference is very large, the clustering quality is poor, and the outliers caused by the derated power are directly eliminated.

Instead of removing power curtailment data as outliers, our goal is to identify derated power operation states and eliminate outliers in wind turbine operating data that clearly deviate from the main trend, which contains several derated power operation states. The difficulty of the problem lies in the fact that there are data points generated by several turbines operating at several reduced power states within the data set. Not only outliers but also power curtailment data points are far from the power curve. Simply assessing distance from the power curve will not help here. Since there are more than one derated power states in the operational data, we cannot directly determine whether a specific data point belongs to the outliers based on the distance from the power curve. Therefore, in this paper, a method for detecting outliers resulting from derated power operation is proposed.

The method converts the outlier detection problem of the wind turbine with derated power operation into a mixed probability distribution model by introducing reasonable assumptions. The K-means clustering algorithm is used to initialize the parameters of the mixed probability model. Then, the expectation maximization (EM) algorithm is used to

derive the updated expression for the model parameters. The logarithm likelihood function is maximized by an iterative method to obtain the optimal model parameters. Finally, the processing of the outliers of the wind turbine data in the power-reduced state is realized by calculating the posterior probability of the sample. The method proposed in this paper can quickly and efficiently identify the degrees of derating in the operational data, distinguish between normal operational data and several different degrees of derated power data, and eliminate outlier data in each data type to improve the quality of wind turbine operational data. The description of this approach is in four parts. The first section above introduced the research background and research status of the wind derated power operation data outlier detection method. The second section introduces the outlier detection model of the wind turbine derated power operation data. In the third section, real operating data of wind turbines located in North China are used to verify the proposed method. Finally, the conclusions are summarized.

2. Wind Turbine Derated Power Operation Data Outlier Detection Model

2.1. Modeling of Derated Power Operation Data Outlier Detection

The method first identifies derated power levels contained in the operational data and then divides the data accordingly for derated power. Finally, the outliers are removed from the operational data corresponding to each type of derated power state. The notion of derated power levels and outliers are shown as Figure 1.

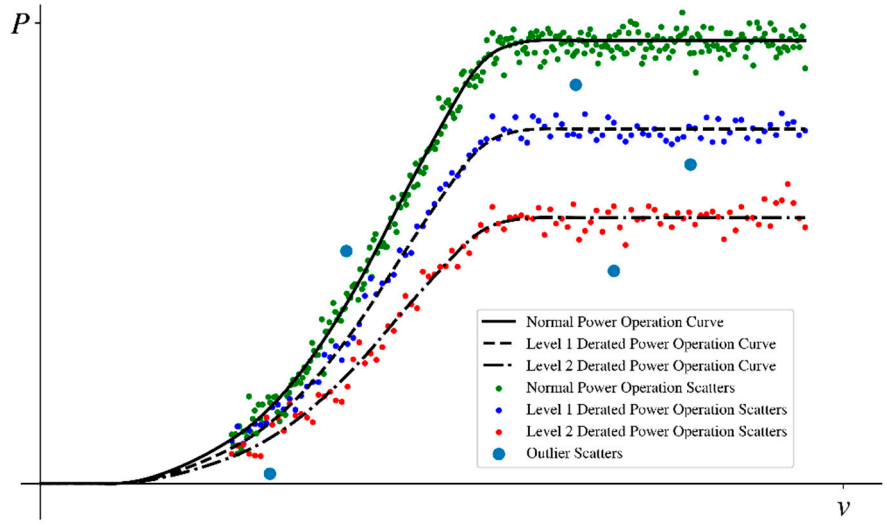


Figure 1. Schematic diagram of derated power operation status and outliers.

In keeping with statistical notation, random variables X, Y, Z represent the wind turbine output power, nacelle wind speed, and derated power state, respectively. Among them, X, Y are observable random variables. Z is a latent random variable and cannot be directly observed from the sample. From the SCADA system of the wind turbine, it is usually easy to observe a sample set of turbine output power and wind speed pairs $\left\{\left(x^{(1)}, y^{(1)}\right), \ldots, \left(x^{(m)}, y^{(m)}\right)\right\}$. Among them, $x^{(i)}$ and $y^{(i)}$ represent the ith output power and cabin wind speed samples in the data set, respectively. We need to establish the posterior probability distribution $p_{Z|XY}(z^{(i)}|x^{(i)}, y^{(i)})$ of the output power X and the nacelle wind speed Y according to the sample data set Z, to identify the power limit of the sample $\left(x^{(i)}, y^{(i)}\right)$, and further remove the outliers from the sample set whose posterior probability value is too low. In order to achieve this goal, a mixed probability density distribution $p_{XY}(x, y)$ with a derated power state Z as a latent random variable is firstly established.

2.2. Mixed Probability Density Distribution Model

In order to establish the mixed probability density distribution model $p_{XY}(x, y)$, the wind turbine derated power assumption and the derated power operation output assumption are introduced.

Power-limited state assumption: the derated power state of the turbine can be expressed in a limited state, that is, the wind turbine derated power state Z can take K different values. They respectively correspond to the normal operating state of the turbine and the $K-1$ derated power operating status with different derated power levels.

Power-limited state output assumption: the output of the wind turbine in a derated power operation state can be expressed as the theoretical output power multiplied by the corresponding derated power degree coefficient. That is, in a derated power state, the output power of the wind turbine can be expressed as $P_k(v) = \alpha_k f(v)$, where $P = f(v)$ is a function of the wind turbine theoretical power curve, and v represents the incoming wind speed. α_k is the power limit coefficient corresponding to the kth derated power state, $\alpha_k \in [0, 1]$, the smaller the value of α_k is, the greater the power limit of the unit is represented. The closer the value of α_k is to 1, the closer the unit state is to the normal power generation state. Here, the normal operating state can be regarded as a special derated power state in which the derated power degree coefficient α_k takes a value of 1.

In order to simplify the modeling process, we use the equal-width discrete method to discretize the wind speed data and divide the wind speed distribution interval into J equal parts. The spacing of each part is the same, and the median value of each wind speed interval represents the interval wind speed. It is further assumed that the discretized wind speed obeys the multinoulli distribution, i.e., $Y \sim \text{Multinoulli}(\psi)$, where vector ψ is the distribution parameter of the polynomial distribution and the jth element of the vector ψ satisfies $\psi_j \geq 0$, $\sum_{j=1}^{J} \psi_j = 1$ and $p_Y(j) = \psi_j$. It can be seen that the probability is $p_Y\left(y^{(i)}\right) = \sum_{j=1}^{J} I\left\{y^{(i)} \in V_j\right\} \cdot \psi_j$, when the wind speed is $y^{(i)}$, where $I\{\cdot\}$ represents the indication function. If the expression in the braces is true, the function value is 1; otherwise, the function value is 0; $I\left\{y^{(i)} \in v_j\right\}$ indicates whether the wind speed value $y^{(i)}$ corresponding to the sample i belongs to the discretized jth wind speed interval V_j.

The turbine's derated power state Z cannot be directly observed; it is thus a latent random variable. It is assumed that the power-limited state also obeys the multinoulli distributions, that is, $Z \sim \text{Multinoulli}(\phi)$, where the kth element of the vector ϕ satisfies $\phi_k \geq 0$, $\sum_{k=1}^{m} \phi_k = 1$ and $p_Z(k) = \phi_k$.

Assume that under a given wind speed and derated power state, the turbine output power obeys a Gaussian distribution, i.e., $X|Y = y^{(i)}, Z = k \sim N(\mu_k\left(y_{(i)}\right), \sigma_k\left(y^{(i)}\right)$, where $\mu_k\left(y_{(i)}\right)$ and $\sigma_k\left(y^{(i)}\right)$ represent the mean and standard deviation of the Gaussian distribution at a given wind speed $y^{(i)}$ and a derated power k, respectively. According to the assumption of the derated power operation output of the wind turbine, the mean of the Gaussian distribution can be expressed as $\mu_k\left(y^{(i)}\right) = \alpha_k f\left(y^{(i)}\right)$; the standard deviation $\sigma_k\left(y^{(i)}\right) = \sum_{j}^{J} I\left\{y^{(i)} \in V_j\right\} \cdot \sigma_{jk}$, where α_k is the derated power coefficient corresponding to the kth derated power state, and σ_{jk} is the standard deviation of the wind speed $y^{(i)}$ in the wind speed interval V_j, and the derated power degree is taken as k.

Furthermore, the wind speed Y and the derated power state Z are independent of each other; thus, $p_{Y|Z}(y|z) = p_Y(y)$.

According to the above assumptions, the mixed probability density distribution $p_{XY}(x,y)$ can be expressed as follows according to the conditional probability and the full probability formula:

$$\begin{aligned} p_{XY}\left(x^{(i)},y^{(i)}\right) &= \sum_{z^{(i)}} p_{XYZ}\left(x^{(i)},y^{(i)},z^{(i)}\right) \\ &= \sum_{z^{(i)}} p_{X|YZ}\left(x^{(i)}|y^{(i)},z^{(i)}\right) p_{Y|Z}\left(y^{(i)}|z^{(i)}\right) p_Z\left(z^{(i)}\right) \\ &= \sum_{z^{(i)}} p_{X|YZ}\left(x^{(i)}|y^{(i)},z^{(i)}\right) p_Y\left(y^{(i)}\right) p_Z\left(z^{(i)}\right) \end{aligned} \tag{1}$$

Among them, $p_{X|YZ}\left(x^{(i)}|y^{(i)},z^{(i)}\right)$ indicates that under the condition that the power-limited state Z is $z^{(i)}$ and the wind speed random variable Y takes $y^{(i)}$, the conditional probability of the unit output power X can be calculated by Equation (2).

$$p_{X|YZ}\left(x^{(i)}|y^{(i)},k\right) = \frac{1}{\sqrt{2\pi}\sigma_k\left(y^{(i)}\right)} \exp\left(-\frac{1}{2\sigma_k^2\left(y^{(i)}\right)}\left(x^{(i)}-\alpha_k f\left(y^{(i)}\right)\right)^2\right) \tag{2}$$

Figure 2 shows the joint probability distribution of wind speed and power in a derated power operation state of a wind turbine in this paper. If the above distribution function parameters $\alpha_k, \sigma_{jk}, \psi_j, \phi_k$ are obtained, the probability values of the items in Equation (1) can be obtained. As shown in Equation (3), we can calculate the posterior probability $p_{Z|XY}(k|x^{(i)},y^{(i)})$ of each sample i under each derated power state k according to the Bayesian formula and then calculate the power-limited state to which the sample i belongs $c^{(i)} = \underset{k}{\operatorname{argmax}}\ p_{Z|XY}(k|x^{(i)},y^{(i)})$. Finally, for each power-limited state k, the set of samples whose posterior probability is lower than the threshold $\left\{(x^{(i)},y^{(i)})\,|c^{(i)}=k, p_{Z|XY}\left(k|x^{(i)},y^{(i)}\right)<\theta\right\}$ is marked as outlier data to achieve outlier detection of wind turbine derated power operation data.

$$p_{Z|XY}(k|x^{(i)},y^{(i)}) = \frac{p_{X|YZ}\left(x^{(i)}|y^{(i)},k;\alpha,\sigma\right)p_Y\left(y^{(i)};\psi\right)p_Z(k;\phi)}{\sum_{k=1}^{K} p_{X|YZ}\left(x^{(i)}|y^{(i)},k;\alpha,\sigma\right)p_Y\left(y^{(i)};\psi\right)p_Z(k;\phi)} \tag{3}$$

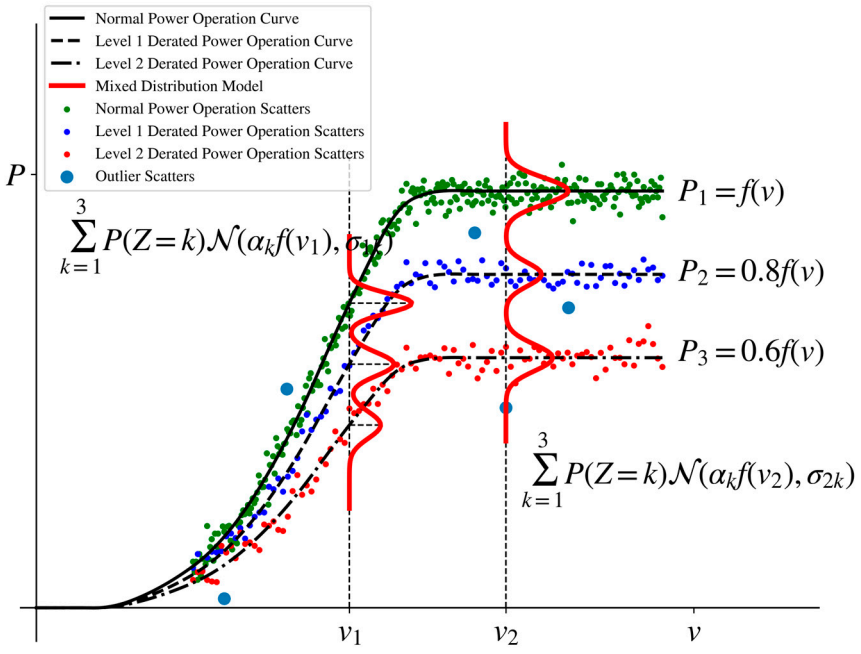


Figure 2. Schematic diagram of the mixed probability distribution model.

To estimate the model parameters, a log-likelihood function can be written, as shown in Equation (4). Since Z is a latent random variable and cannot be directly observed, it is

difficult to directly maximize the log-likelihood function (3) to solve the parameters. We turn to the idea of the EM algorithm to solve the problem.

$$\mathcal{L}(\alpha,\sigma,\psi,\phi)=\sum_i \log p_{XY}\left(x^{(i)},y^{(i)};\alpha,\sigma,\psi,\phi\right) \tag{4}$$

2.3. EM Algorithm Estimation Model Parameters

According to the idea of the EM algorithm, we do not directly solve the maximum value of the log-likelihood function, and instead go to the lower bound (E-step) of the log-likelihood function and then maximize the lower bound (M-step). We can find the model parameters that maximize the likelihood function by iteratively repeating the E-step and M-step. We firstly introduce the Jensen inequality.

Theorem. Let f be a convex function and let X be a random variable.

Then:

$$E[f(X)]\geq f(E[X]) \tag{5}$$

Moreover, if f is strictly convex, then $E[f(X)]=f(E[X])$ holds true if and only if $X=E[X]$ with probability 1 (i.e., if X is a constant).

According to Jensen's inequality, the lower bound of the log-likelihood function can be obtained:

$$\begin{aligned}\mathcal{L}(\alpha,\sigma,\psi,\phi) &=\sum_i \log p_{XY}\left(x^{(i)},y^{(i)};\alpha,\sigma,\psi,\phi\right)\\ &=\sum_i \log \sum_{z^{(i)}} p_{XYZ}\left(x^{(i)},y^{(i)},z^{(i)}\right)\\ &=\sum_i \log \sum_{z^{(i)}} Q_i\left(z^{(i)}\right)\frac{p_{XYZ}\left(x^{(i)},y^{(i)},z^{(i)}\right)}{Q_i\left(z^{(i)}\right)}\\ &\geq \sum_i \sum_{z^{(i)}} Q_i\left(z^{(i)}\right)\log\frac{p_{XYZ}\left(x^{(i)},y^{(i)},z^{(i)}\right)}{Q_i\left(z^{(i)}\right)}\end{aligned} \tag{6}$$

Among them, $\mathcal{L}(\alpha,\sigma,\psi,\phi)$ is a log-likelihood function of the mixed probability model; Q represents a certain distribution, and the condition that the inequality takes the equal sign is that $\frac{p_{XYZ}\left(x^{(i)},y^{(i)},z^{(i)}\right)}{Q_i\left(z^{(i)}\right)}$ is a constant. According to $\sum_{z^{(i)}} Q_i\left(z^{(i)}\right)=1$, you can obtain:

$$\begin{aligned}Q_i\left(z^{(i)}\right) &=\frac{p_{XYZ}\left(x^{(i)},y^{(i)},z^{(i)}\right)}{\sum_{z^{(i)}} p_{XYZ}\left(x^{(i)},y^{(i)},z^{(i)}\right)}\\ &=p_{Z|XY}\left(z^{(i)}|x^{(i)},y^{(i)}\right)\end{aligned} \tag{7}$$

Let $w_k^{(i)}=Q_i(k)=p_{Z|XY}(k|x^{(i)},y^{(i)})$. According to the Bayesian formula, you can obtain:

$$w_k^{(i)}=\frac{p_{X|YZ}\left(x^{(i)}|y^{(i)},k;\alpha,\sigma\right)p_Y\left(y^{(i)};\psi\right)p_Z(k;\phi)}{\sum_{k=1}^{K} p_{X|YZ}\left(x^{(i)}|y^{(i)},k;\alpha,\sigma\right)p_Y\left(y^{(i)};\psi\right)p_Z(k;\phi)} \tag{8}$$

Let $\ell(\alpha,\sigma,\psi,\phi)$ take the right side of the inequality of Equation (5) as the lower bound of the log-likelihood function, then $\ell(\alpha,\sigma,\psi,\phi)$ can be expressed as:

$$\ell(\alpha,\sigma,\psi,\phi)=\sum_{i=1}^{m}\sum_{k=1}^{K} w_k^{(i)}\log\frac{\frac{1}{\sqrt{2\pi}\sigma_k\left(y^{(i)}\right)}\exp[-\frac{1}{2\sigma_k^2\left(y^{(i)}\right)}\left(x^{(i)}-\alpha_k f\left(y^{(i)}\right)^2\right]\psi\left(y^{(i)}\right)\phi_k}{w_k^{(i)}} \tag{9}$$

where $\psi\left(y^{(i)}\right)=p_Y\left(y^{(i)}\right)=\sum_{j=1}^{J} I\left(y^{(i)}\in V_j\right)\cdot\psi_j$, $\sigma_k\left(y^{(i)}\right)=\sum_{j=1}^{J} I\left(y^{(i)}\in V_j\right)\cdot\sigma_{jk}$.

After obtaining the lower bound $\ell(\alpha,\sigma,\psi,\phi)$ of the log-likelihood function, we can obtain the partial derivative of the lower bound on the parameters α,σ,ψ,ϕ. Then, let the partial derivative equal zero, and obtain the model parameters by maximizing the lower bound $\ell(\alpha,\sigma,\psi,\phi)$ of the log-likelihood function.

Find the partial derivative of ℓ to α_q

$$\frac{\partial \ell}{\partial \alpha_q} = \sum_{i=1}^{m} w_q^{(i)} \frac{[x^{(i)} - \alpha_q f\left(y^{(i)}\right)] f\left(y^{(i)}\right)}{\sigma_q\left(y^{(i)}\right)} \tag{10}$$

Let the above formula be equal to zero, and we find:

$$\alpha_q = \frac{\sum_{i=1}^{m} w_q^{(i)} x^{(i)} f\left(y^{(i)}\right)}{\sum_{i=1}^{m} w_q^{(i)} f^2\left(y^{(i)}\right)} \tag{11}$$

Next, we find the partial derivative of ℓ on σ_{pq}:

$$\begin{aligned} \frac{\partial \ell}{\partial \sigma_{pq}} &= \sum_{i=1}^{m} \frac{\partial}{\partial \sigma_{pq}} w_q^{(i)} I\left(y^{(i)} \in V_p\right) [-\log \sigma_{pq} - \frac{1}{2\sigma_{pq}^2}\left(x^{(i)} - \alpha_q f\left(y^{(i)}\right)\right)^2] \\ &= \sum_{i}^{m} w_q^{(i)} I\left(y^{(i)} \in V_p\right) \left[-\frac{1}{\sigma_{pq}} + \frac{1}{\sigma_{pq}^3}\left(x^{(i)} - \alpha_q f\left(y^{(i)}\right)\right)^2\right] \end{aligned} \tag{12}$$

Let the above formula be equal to zero, and we find:

$$\sigma_{pq}^2 = \frac{\sum_{i=1}^{m} w_q^{(i)} I\left(y^{(i)} \in V_p\right) \left(x^{(i)} - \alpha_q f\left(y^{(i)}\right)\right)^2}{\sum_{i=1}^{m} w_q^{(i)} I\left(y^{(i)} = p\right)} \tag{13}$$

Find the partial derivative of ℓ on ϕ_q. $\sum_{k=1}^{K} \phi_k = 1$. Using the Lagrangian multiplier method, we find the partial derivative of $\ell + \lambda\left(\sum_{k=1}^{K} \phi_k - 1\right)$ on ϕ_q and λ.

$$\frac{\partial}{\partial \phi_q}\left(\ell + \lambda\left(\sum_{k=1}^{K} \phi_k - 1\right)\right) = \sum_{i}^{m} \frac{w_q^{(i)}}{\phi_q} + \lambda \tag{14}$$

$$\frac{\partial}{\partial \lambda}\left(\ell + \lambda\left(\sum_{k=1}^{K} \phi_k - 1\right)\right) = \sum_{k=1}^{K} \phi_k - 1 \tag{15}$$

Let the upper two formulas be equal to zero. Combine these two formulas and we can obtain the solution:

$$\phi_q = \frac{1}{m} \sum_{i=1}^{m} w_q^{(i)} \tag{16}$$

In the same way, $\sum_{j=1}^{J} \psi_j = 1$. Using the Lagrangian multiplier method, we find the partial derivative of $\ell + \gamma\left(\sum_{j=1}^{J} \psi_j - 1\right)$ on ψ_p and γ.

$$\begin{aligned} \frac{\partial}{\partial \psi_p}\left(\ell + \gamma\left(\sum_{j}^{J} \psi_j - 1\right)\right) &= \frac{\partial}{\partial \psi_p}\left(\sum_{i=1}^{m} \sum_{k=1}^{K} w_k^{(i)} I\left(y^{(i)} \in V_p\right) \cdot \log \psi_p + \gamma\left(\sum_{j=1}^{J} \psi_j - 1\right)\right) \\ &= \sum_{i=1}^{m} \sum_{k=1}^{K} \frac{w_k^{(i)} I\left(y^{(i)} \in V_p\right)}{\psi_p} + \lambda \end{aligned} \tag{17}$$

$$\frac{\partial}{\partial \gamma}\left(\ell + \gamma\left(\sum_{j}^{J} \psi_j - 1\right)\right) = \sum_{j}^{J} \psi_j - 1 \tag{18}$$

Let the upper two formulas be equal to zero. Combine these two formulas and we can obtain the solution:

$$\psi_p = \frac{\sum_{i=1}^{m} \sum_{k=1}^{K} I\left(y^{(i)} \in V_p\right) w_k^{(i)}}{m} \tag{19}$$

Thus far, we have derived two main processes, E-step and M-step, in the EM algorithm. In the E-step, according to the initialized parameters, we can calculate $w_k^{(i)}$ according to Formula (8). In the M-step, the parameters $\alpha, \sigma, \psi, \phi$ are updated according to Equations (11), (13), (16) and (19); the E-step and M-step are repeated repeatedly until convergence. Then, we can find the model parameters.

Convergence is guaranteed by the EM algorithm. Hence, we will no longer discuss the proof process in this paper. However, the EM algorithm can only converge to the local optimum, and the result is greatly affected by the initial value. Here, we will give the identification method of the optimal derated power level number K and the initialization method of other parameters of the model to help the EM algorithm to quickly and stably converge.

2.4. Optimal Derated Power Level Identification

The parameters of the wind turbine derated power operation data anomaly detection model proposed in this paper include: the number of derated power levels K; the number of discretized wind speed intervals J; the derated power coefficient $\alpha = [\alpha_1, \ldots, \alpha_K]$, the power limit distribution parameter $\phi = [\phi_1, \ldots, \phi_K]$, the discretized wind speed probability distribution parameter $\psi = [\psi_1, \ldots, \psi_J]$ and the mean parameter of the Gaussian distribution μ_{jk} and the variance parameter σ_{jk}^2. We firstly determine the optimal derated power level number K.

Inspired by the k-means clustering algorithm, we firstly give the derated power level number K and the discretized wind speed interval number J and random initialization $\alpha = [\alpha_1, \ldots, \alpha_K]$, $\psi = [\psi_1, \ldots, \psi_K]$, $\phi = [\phi_1, \ldots, \phi_K]$, μ_{jk} and σ_{jk}^2. For each sample i, we calculate the distance $d_k^{(i)} = |x^{(i)} - \alpha_k f\left(y^{(i)}\right)|$ from $\left(x^{(i)}, y^{(i)}\right)$ to each of the derated power output curves. We find the derated power state corresponding to the curve with the smallest distance from the sample i to each of the derated power state running curves as the power limit state of the sample $\left(x^{(i)}, y^{(i)}\right)$, denoted as $c^{(i)} = \underset{k}{\operatorname{argmin}}\, d_k^{(i)}$. The distance of the sample i from the corresponding derated power running curve is recorded as $d^{(i)} = \min\limits_{k} d_k^{(i)}$. For the sample set of the same power state $\left\{ \left(x^{(i)}, y^{(i)}\right) | c^{(i)} = k \right\}$, we use the least squares fitting derated power operation curve $x = \alpha_k f(y)$ and update the corresponding derated power coefficient α_k.

Let K take 2-8 in sequence and repeat the above steps several times. We calculate the average loss value of all samples $\sum\limits_{i} \frac{d^i}{m}$ and take the average value of each loss. We use the mean as the vertical axis and the K value as the horizontal axis as the elbow curve. The K value corresponding to the position where the average loss function value has the largest decrease is taken as the optimal derated power level number.

2.5. Model Parameter Initialization Method

After obtaining the optimal power level number K and the power limit class $c^{(i)}$ corresponding to each sample, we can directly write the initialization expression of the parameters ψ, ϕ, μ, σ, as shown in Formula (20) to Formula (23).

$$\psi_j = \frac{\sum_{i=1}^{m} I\left\{y^{(i)} \in V_j\right\}}{m} \tag{20}$$

$$\phi_k = \frac{\sum_{i=1}^{m} I\left\{c^{(i)} = k\right\}}{m} \tag{21}$$

$$\mu_{jk} = \frac{\sum_{i=1}^{m} I\left\{c^{(i)} = k, y^{(i)} \in V_j\right\} x^{(i)}}{m} \tag{22}$$

$$\sigma_{jk}^2 = \frac{\sum_{i=1}^{m} I\left\{c^{(i)} = k, y^{(i)} \in V_j\right\} \left(x^{(i)} - \mu_{jk}\right)^2}{m} \tag{23}$$

According to the derated power level $c^{(i)}$ corresponding to each sample, we can initially realize the division of the derated power data, but the result obtained by this process is not the optimal result, and based on the k-means clustering algorithm, it is also impossible to eliminate outliers in each of the derated power levels. However, we can use the parameters obtained in the above process as the initial values of the EM algorithm, so that the parameters of the final model can be stably converged to similar local best points.

Based on the above modeling process and parameter initialization process, the algorithm for the anomaly detection method of wind turbine derated power operation data proposed in this paper is as follows (Algorithm 1):

Algorithm 1 Wind turbine derated power operation data outlier detection

Require: the sample set of turbine output power and nacelle wind speed pair $\left\{\left(x^{(1)}, y^{(1)}\right), \ldots, \left(x^{(m)}, y^{(m)}\right)\right\}$, discretized wind speed interval number $J = 50$, wind turbine theoretical power curve function $P = f(v)$, probability threshold θ.

1. Initialize the model parameters to obtain the optimal power level number K and the initial values of the mixed probability model parameters $\alpha^{(0)}, \sigma^{(0)}, \psi^{(0)}, \phi^{(0)}$.
2. While $\alpha, \sigma, \psi, \phi$ have no convergence. Do
3. E-step: For each sample i and the derated power k, calculate $w_k^{(i)} = \frac{p_{X|YZ}\left(x^{(i)}|y^{(i)},k;\alpha,\sigma\right)p_Y\left(y^{(i)};\psi\right)p_Z(k;\phi)}{\sum_{k=1}^{K} p_{X|YZ}\left(x^{(i)}|y^{(i)},k;\alpha,\sigma\right)p_Y\left(y^{(i)};\psi\right)p_Z(k;\phi)}$.
4. M-step: Update parameters
5. $\alpha_q = \frac{\sum_{i=1}^{m} w_q^{(i)} x^{(i)} f\left(y^{(i)}\right)}{\sum_{i=1}^{m} w_q^{(i)} f^2\left(y^{(i)}\right)}$
6. $\sigma_{pq}^2 = \frac{\sum_{i=1}^{m} w_q^{(i)} I\left(y^{(i)} \in V_p\right)\left(x^{(i)} - \alpha_q f\left(y^{(i)}\right)^2\right.}{\sum_{i=1}^{m} w_q^{(i)} I\left(y^{(i)} \in V_p\right)}$
7. $\phi_q = \frac{1}{m}\sum_{i=1}^{m} w_q^{(i)}$
8. $\psi_p = \frac{\sum_{i=1}^{m}\sum_{k=1}^{K} I\left(y^{(i)} \in V_p\right) w_k^{(i)}}{m}$
9. End while.
10. Calculate the posterior probability $p_{Z|XY}(k|x^{(i)}, y^{(i)})$ of each sample i under each derated power state k according to iterative parameters.
11. Calculate the derated power level $c^{(i)} = \underset{k}{\operatorname{argmax}}\, p_{Z|XY}(k|x^{(i)}, y^{(i)})$ to which the sample i belongs.
12. For each derated power k, the sample set whose posterior probability is lower than the threshold $\left\{(x^{(i)}, y^{(i)}) \,|c^{(i)} = k, p_{Z|XY}\left(k|x^{(i)}, y^{(i)}\right) < \theta\right\}$ is marked as outlier data.

3. Analysis of the Case of Outlier Detection of Wind Power Turbine Derated Power Operation Data

SCADA data of a wind farm in North China are used to verify the effectiveness of the proposed outlier detection method for wind turbine derated power operation data. North China is a location where "wind curtailment" is common. Wind data from this area are suitable for studying data processing methods for derated power operation. In order to fully verify the method proposed in this paper, we select the operating data of three 2.5 MW direct-drive permanent-magnet synchronous generator wind turbines (#1, #2, #6) with different derated power levels on the wind farm. The basic information for these three wind turbines is shown in Table 1. The operating data of the SCADA system of the turbine were intercepted from 10:30 a.m. on 20 November 2014 to 10:10 a.m. on 19 January 2015 for analysis.

3.1. Case Analysis

The method proposed in this paper will be described in a detailed way by taking the #2 turbine as an example. We plot the wind speed and power scatter plot of the turbine and the manufacturer's power curve that corrected for air density in the standard manner, as shown in Figure 3. It can be seen from the figure that the power scatter plot of turbine #2 is approximately distributed over three clusters and that a large proportion of the points reflect derated power operation. The uppermost cluster of scatter points is closest to the manufacturer's power curve, and the turbine can be considered here to be in normal operation. The following two clusters of scatter points correspond to different derated power states of the turbine. There are still some outliers in the scatter plot that deviate significantly from the main trends, which we have to identify and eliminate. Since there is more than one derated power state in the operational data, we cannot directly determine

whether a specific data point belongs to the outliers based on the distance from the power curve. To do this, we need to identify and classify the derated power operating states of the operational data and then remove those outliers from the samples corresponding to each type of derated power state.

Table 1. Basic parameters of wind turbines.

Parameter	Value
Rated power/kW	2500
Power adjustment method	Variable pitch, variable speed
Number of blades	3
Rotor diameter/m	103
Hub height/m	80
Cut-in wind speed/(m/s)	3
Rated wind speed/(m/s)	11
Cut-out wind speed/(m/s)	25
Maximum wind speed/(m/s)	39.6
Extreme wind speed/(m/s)	55
Air density/kg/m^3	1.0622

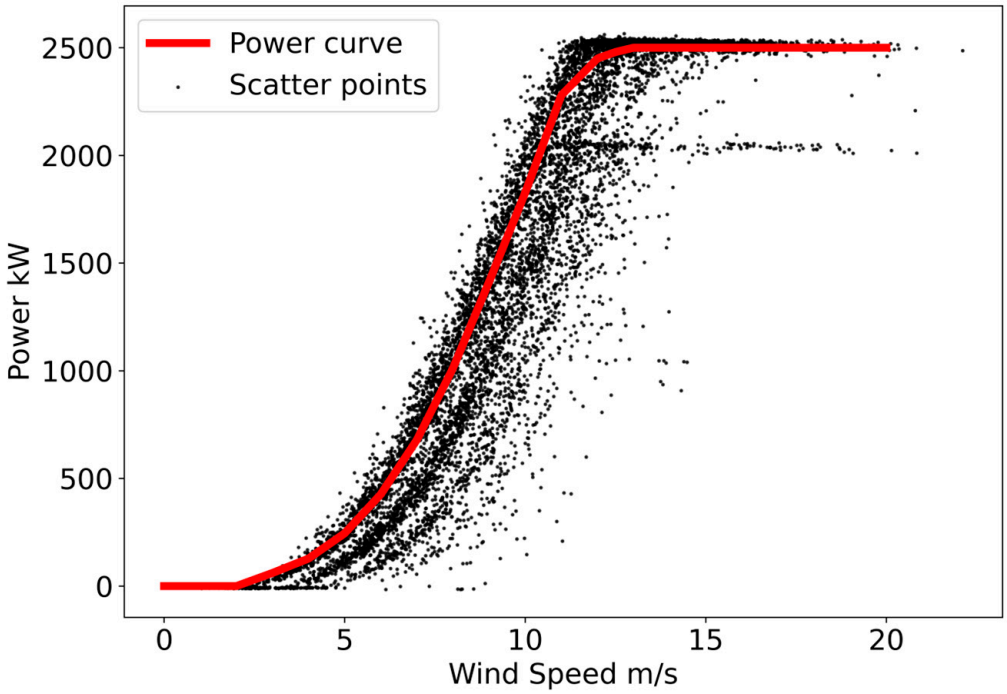


Figure 3. Wind speed and power scatter plot of turbine #2 and the manufacturer's power curve that corrected for air density in the standard manner.

In order to determine the optimal derated power level, the elbow curve is drawn according to the best derated power level identification method proposed in Section 2.4, as shown in Figure 4. It can be seen from the figure that the K value corresponding to the position where the average loss function value has the largest decrease is 3. It can be seen that this set of data contains one set of normal operating states and two different levels of derated power operating states, which is consistent with the results we observed directly from the data.

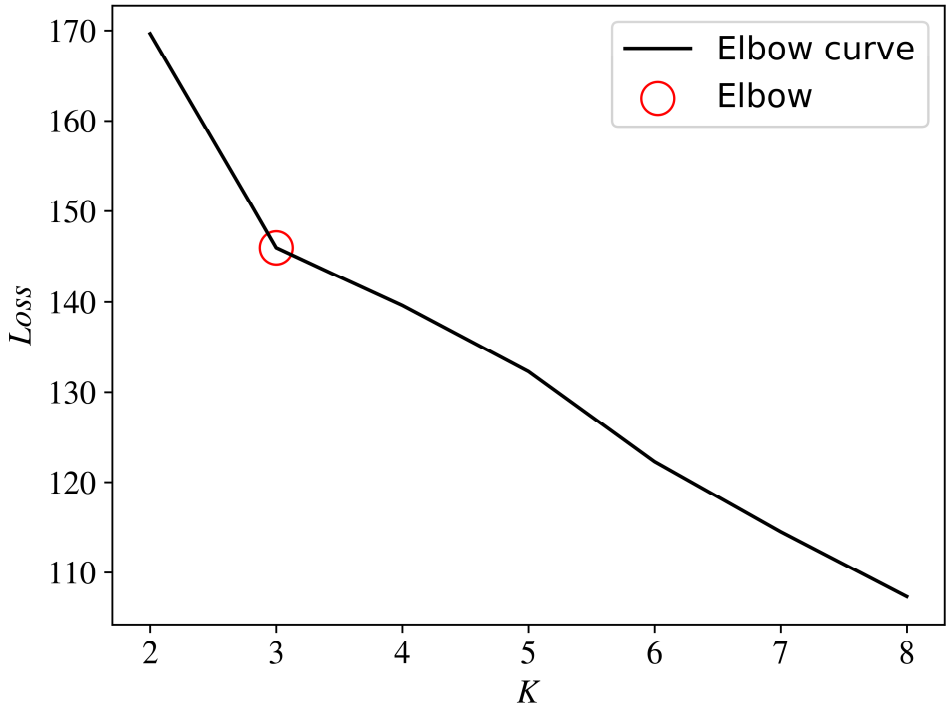


Figure 4. Elbow curve to determine the optimal derated power level.

After the initialization step, we can obtain the preliminary operation data derated power state allocation, as shown in Figure 5. The result is obtained based on the distance of the scatter distance from each of the derated power curves. It can be seen from the figure that the scatter points of red, green and blue respectively represent the division of the normal operating state and the provisional first and second derated power levels given by the model. This step preliminarily realizes the division of the derated power level of the operational data, but it is incomplete. We will initialize the parameters of the probability model based on this result to achieve a more refined model.

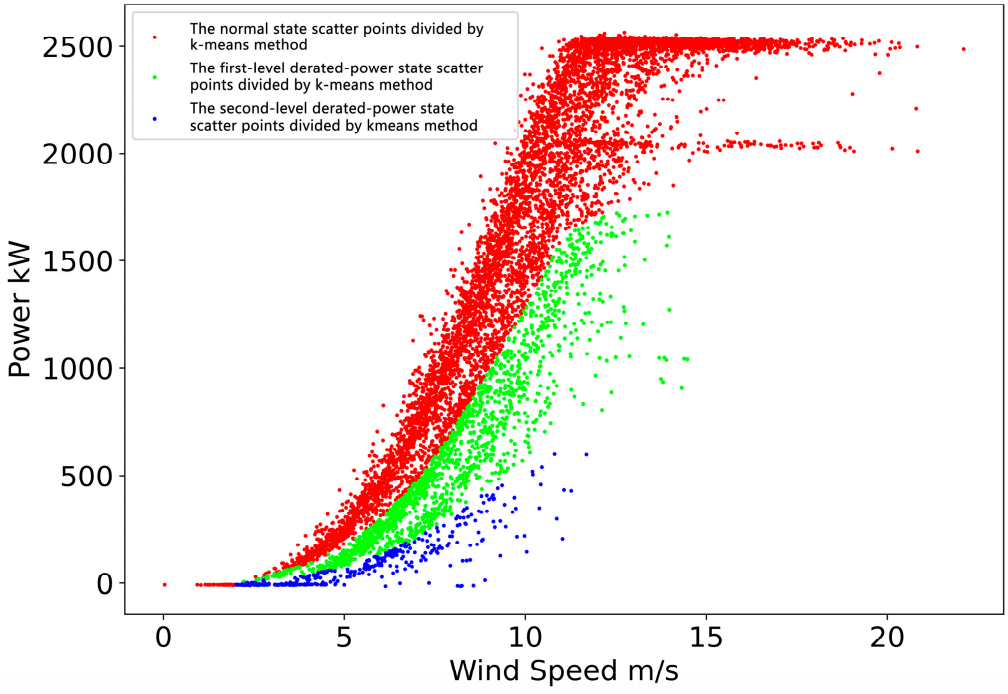


Figure 5. Operation data derated power state division result obtained by k-means method.

According to the method proposed in Section 2.2, the mixed probability model is established, and the mixed probability model parameters are estimated by the EM algorithm according to the proposed method in Section 2.3. Then, the posterior probability $p_{Z|XY}(k|x^{(i)}, y^{(i)})$ of each sample i under each derated power state k is calculated according to the Bayesian formula. Then, the probability value $c^{(i)} = \underset{k}{\operatorname{argmax}}\, p_{Z|XY}(k|x^{(i)}, y^{(i)})$ of the derated power level to which the sample i belongs is calculated to implement the division of the derated power state of the operating data. Figure 6 shows the results of the partitioning of the mixed probability model. We use different colors to represent the derated power state to which each sample belongs. The red points represent a higher probability that the sample is in normal operation. The green and blue points respectively represent a higher probability that the sample belongs to the derated power state 1 and the derated power state 2. Sample color between two derated power states will mix. This is because the probability that the sample belongs to these two power-limited states is very close. This reflects the uncertainty of the derated power state to which the sample belongs.

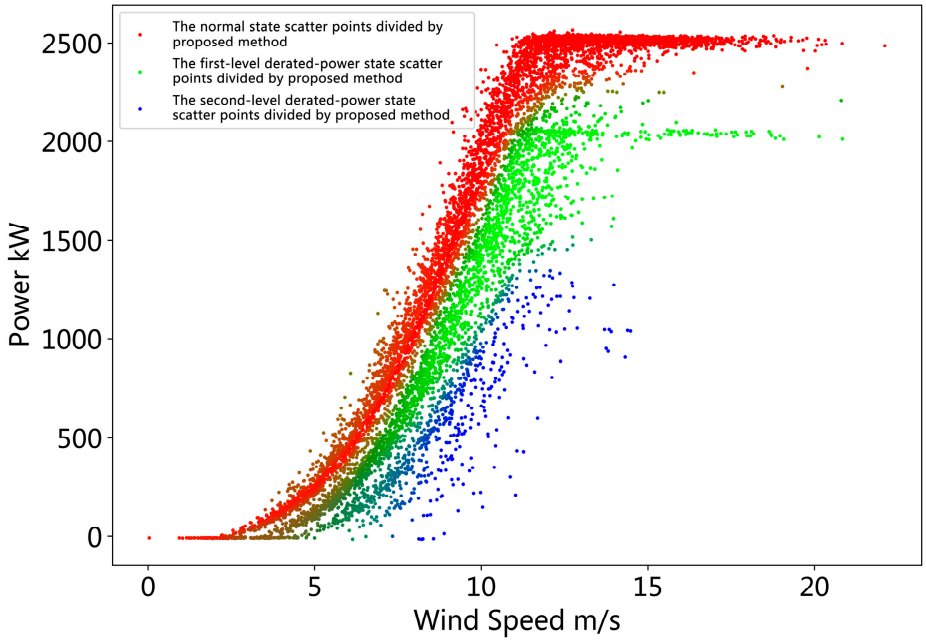


Figure 6. Operational data derated power state division result obtained by the proposed method.

For each derated power state k, we give a threshold $\theta = 0.8$. The sample set $\left\{(x^{(i)}, y^{(i)}) \,|c^{(i)} = k, p_{Z|XY}\left(k|x^{(i)}, y^{(i)}\right) < \theta\right\}$ with the posterior probability below the threshold is eliminated to achieve outlier elimination. The result is shown in Figure 7. Compared with the results of k-means partitioning, the proposed method is more effective. In addition, it will be seen from the following cases that more flexible outlier detection can be achieved by giving different thresholds θ in the process of eliminating outliers.

3.2. Comparison of Outlier Detection Results of Wind Turbine Operation Data with Different Derated Power Levels

In order to further verify the effectiveness of the method, we performed the same steps as those in Section 3.1 for the operation data of turbine #1 and turbine #6. We obtained the derated power data degree division and outlier detection results of the operational data of the two turbines, as shown in Figures 8 and 9. It can be seen from Figure 8a that the method proposed in this paper identifies that the operating data of turbine #1 contains five operating states, including normal operating conditions and four derated power operating states. From the results of the derated power state division, the derated power data are clearly separated from the normal state data, and the four derated power states also achieve

good distinction. After the threshold value θ is set, the outliers corresponding to the respective operating states can be detected. Since there are less derated power data in the operation data of turbine #1, we will not display the operation data under each derated power state and only draw the results after removing the outliers in normal operation, as shown in Figure 8b.

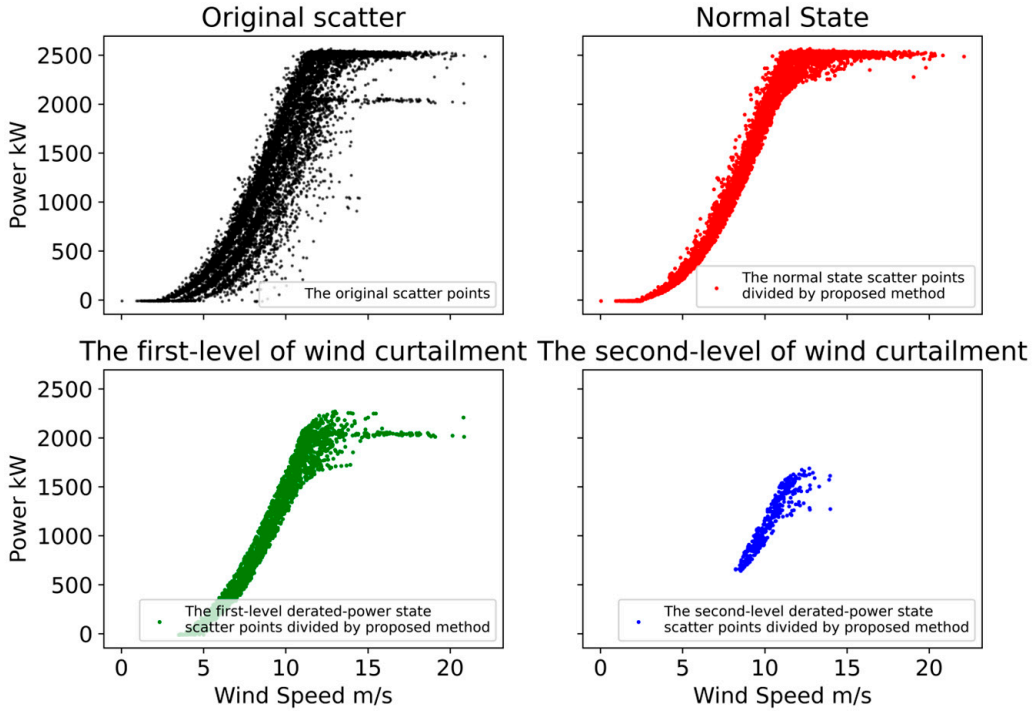


Figure 7. Operation data derated power state division and outliers elimination result.

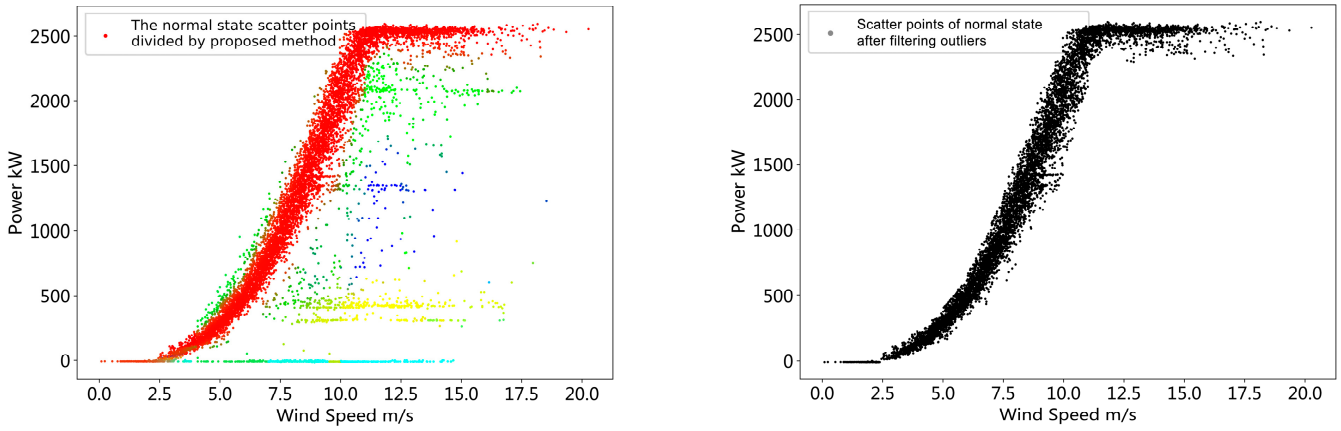
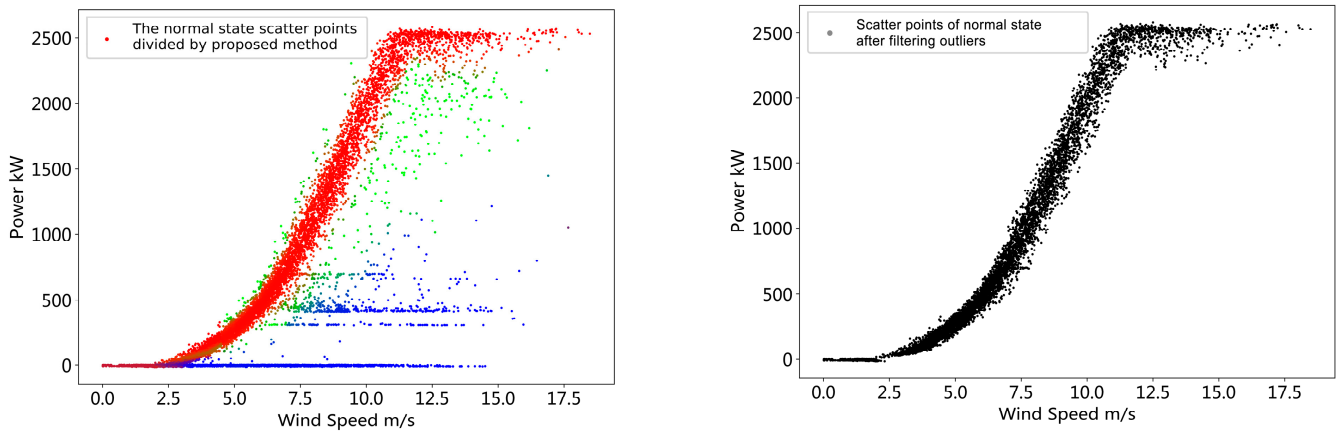


Figure 8. Operation data derated power state division and outlier elimination result of turbine #1. (**a**) The derated power state division result of turbine #1, in this case, five different colors represent five derated power levels in the dataset of turbine #1, and red points represents normal power operational state; (**b**) the outliers elimination result under normal operating conditions of turbine #1.



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

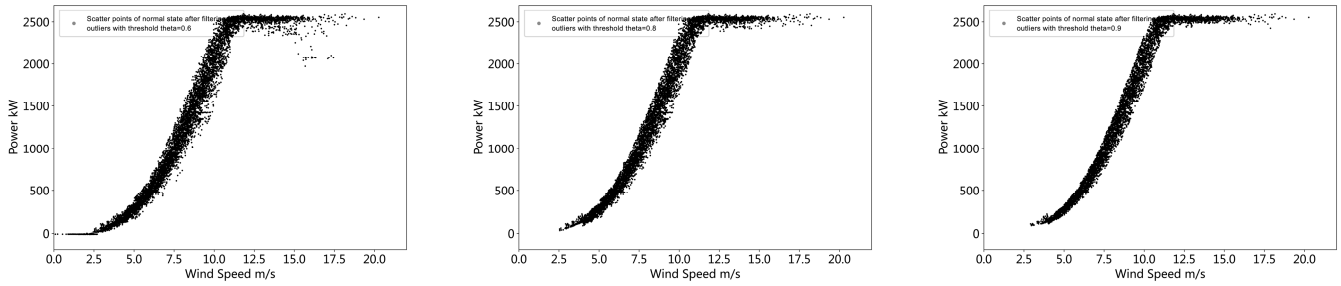
Figure 9. Operation data derated power state division and outlier elimination result of turbine #6. (**a**) The derated power state division result of turbine #6, in this case, three different colors represent three derated power levels in the dataset of turbine #6, and red points represents normal power operational state; (**b**) the outliers elimination result under normal operating conditions of turbine #6.

Similarly, as can be seen from Figure 9a, the method proposed in this paper identifies that the operating data of turbine #6 contains three operating states, including normal operating conditions and two derated power operating states. From the results of the derated power state division, the derated power data and the normal state data achieve a good separation, but the identification of the two derated power states does not achieve the expected effect. The green data points are outliers from normal operation, while the blue data points contain two derated power states. However, this does not affect the outlier detection under normal operating conditions. It can be seen from Figure 9b that the good outlier detection effect is also achieved under the normal operation of turbine #6.

Comparing turbine #1 and turbine #6, we can see that the method proposed in this paper can adapt to different operating numbers of derated power levels and is more reliable in the detection of outliers.

3.3. Impact of Threshold on Outlier Detection Results

The method proposed in this paper relies on the posterior probability value of the sample to detect the outliers. When the posterior probability of the sample is less than the given threshold, the sample is judged as an outlier. Therefore, if different thresholds are applied, different outliers will be identified. To investigate the effect of thresholds on outlier detection, the same turbine data are used with the same model but with different thresholds. We establish a mixed probability model for the operating data of turbine #1 and estimate the model parameters to obtain the posterior probability. The threshold values θ are set to 0.6, 0.8, and 0.9. The outlier detection results are shown in Figure 9. It can be observed from the figure that when the threshold is low (θ = 0.6), more sample data will be retained, and when the threshold is higher (θ = 0.9), less data will be retained. As the threshold gradually increases, the points in the lower probability density regions will be gradually removed, which is consistent with expectations. Therefore, the threshold can be flexibly selected according to the quality requirements of the operational data. Selecting a higher threshold results in fewer samples with outliers, and selecting a lower threshold preserves more raw data. In this case, selecting a threshold of 0.8 eliminates most outliers and preserves as much raw data as possible. Furthermore, from the comparison of Figure 10a–c, it can be found that changing the threshold θ over a larger range does not produce a huge change in the outcome. The result is still able to maintain the data that need to be retained, that is, the model is robust to the choice of threshold.



(**a**) $\theta = 0.6$ (**b**) $\theta = 0.8$ (**c**) $\theta = 0.9$

Figure 10. Impact of threshold θ on ourlier detection result.

4. Conclusions

In this paper, we propose an outlier detection method suitable for wind turbines with derated power operation. The following conclusions can be drawn from the analysis of the example:

1. By introducing reasonable assumptions regarding derated power operation, a mixed probability distribution model of wind speed and power with derated power as an implicit random variable is established. Outlier detection under the derated power level of the wind turbine is realized. The method is based on statistical theory and the expectation maximization algorithm and has a sound theoretical basis.
2. A mixed probability model initialization method based on k-means clustering is established. This method can identify derated power levels in the data and provide the initial value of the parameter of the mixed probability model. It can also accelerate the convergence of model parameters, making the results more stable.
3. The example shows that the method can reliably detect outliers of the wind turbine under normal operating conditions and under different derated power states. The method can also distinguish different derated power state data and has certain robustness to the selection of model parameters.

Author Contributions: Conceptualization, Y.M.; methodology, Y.M.; software, Y.M.; validation, Y.M. and Z.Y.; formal analysis, Y.M.; investigation, Y.M. and T.T.; resources, Y.L.; data curation, Y.L.; writing—original draft preparation, Y.M.; writing—review and editing, Y.L., Z.Y. and D.I.; writing-language polish, D.I.; visualization, Y.M.; supervision, Y.L. and J.Y.; project administration, Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key Research and Development Program of China, grant number 2019YFE0104800.

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. Wu, Y.; He, F.; Zhou, J.; Wu, C.; Liu, F.; Tao, Y.; Xu, C. Optimal Site Selection for Distributed Wind Power Coupled Hydrogen Storage Project Using a Geographical Information System Based Multi-Criteria Decision-Making Approach: A Case in China. *J. Clean. Prod.* **2021**, *299*, 126905. [CrossRef]
2. Dai, J.; Yang, W.; Cao, J.; Liu, D.; Long, X. Ageing Assessment of a Wind Turbine over Time by Interpreting Wind Farm SCADA Data. *Renew. Energy* **2018**, *116*, 199–208. [CrossRef]
3. Gonzalez, E.; Stephen, B.; Infield, D.; Melero, J.J. Using High-Frequency SCADA Data for Wind Turbine Performance Monitoring: A Sensitivity Study. *Renew. Energy* **2019**, *131*, 841–853. [CrossRef]
4. Liu, Y.; Wu, Z.; Wang, X. Research on Fault Diagnosis of Wind Turbine Based on SCADA Data. *IEEE Access* **2020**, *8*, 185557–185569. [CrossRef]

5. Wu, Z.; Sun, H. Behavior of Chinese Enterprises in Evaluating Wind Power Projects: A Review Based on Survey. *Renew. Sustain. Energy Rev.* **2015**, *43*, 133–142. [CrossRef]
6. Schlechtingen, M.; Santos, I.F. Comparative Analysis of Neural Network and Regression Based Condition Monitoring Approaches for Wind Turbine Fault Detection. *Mech. Syst. Signal Process.* **2011**, *25*, 1849–1875. [CrossRef]
7. Kusiak, A.; Zheng, H.; Song, Z. Models for Monitoring Wind Farm Power. *Renew. Energy* **2009**, *34*, 583–590. [CrossRef]
8. Wang, C.; Gao, H.; Liu, Z.; Fu, Y. A New Outlier Detection Model Using Random Walk on Local Information Graph. *IEEE Access* **2018**, *6*, 75531–75544. [CrossRef]
9. Agyemang, M.; Barker, K.; Alhajj, R. A Comprehensive Survey of Numeric and Symbolic Outlier Mining Techniques. *Intell. Data Anal.* **2006**, *10*, 521–538. [CrossRef]
10. Agrawal, S.; Agrawal, J. Survey on Anomaly Detection Using Data Mining Techniques. In Proceedings of the Knowledge-Based and Intelligent Information & Engineering Systems 19th Annual Conference, Kes-2015, Singapore, 7–9 September 2015; Ding, L., Pang, C., Kew, L.M., Jain, L.C., Howlett, R.J., Eds.; Elsevier Science Bv: Amsterdam, The Netherlands, 2015; Volume 60, pp. 708–713.
11. Angiulli, F.; Fassetti, F. DOLPHIN: An Efficient Algorithm for Mining Distance-Based Outliers in Very Large Datasets. *ACM Trans. Knowl. Discov. Data* **2009**, *3*, 4. [CrossRef]
12. Zemene, E.; Tesfaye, Y.T.; Prati, A.; Pelillo, M. Simultaneous Clustering and Outlier Detection Using Dominant Sets. In Proceedings of the 2016 23rd International Conference on Pattern Recognition (icpr), Cancun, Mexico, 4–8 December 2016; IEEE Computer Society: Los Alamitos, CA, USA, 2016; pp. 2325–2330.
13. Zhu, Q.; Ye, L.; Zhao, Y.; Lang, Y.; Song, X. Methods for elimination and reconstruction of abnormal power data in wind farms. *Power Syst. Prot. Control* **2015**, *3*, 38–45. (In Chinese)
14. Zhang, D.; Li, W.; Liu, Y.; Liu, C. Reconstruction Method of Active Power Historical Operating Data for Wind Farm. *Autom. Electr. Power Syst.* **2021**, *5*, 14–18, 24.
15. Wang, Y.; Infield, D.G.; Stephen, B.; Galloway, S.J. Copula-Based Model for Wind Turbine Power Curve Outlier Rejection: Copula-Based Model for Wind Turbine Power Curve Outlier Rejection. *Wind Energy* **2014**, *17*, 1677–1688. [CrossRef]
16. Wang, Z.; Wang, L.; Huang, C. A Fast Abnormal Data Cleaning Algorithm for Performance Evaluation of Wind Turbine. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 5001309. [CrossRef]
17. Shen, X.J.; Fu, X.J.; Zhou, C.C. Characteristics of Outliers in Wind Speed-Power Operation Data of Wind Turbines and Its Cleaning Method. *Trans. China Electrotech. Soc.* **2020**, *33*, 3353–3361. [CrossRef]
18. Liu, X.; Ren, H.; Wang, G. Computing Halfspace Depth Contours Based on the Idea of a Circular Sequence. *J. Syst. Sci. Complex.* **2015**, *28*, 1399–1411. [CrossRef]
19. Black, I.M.; Richmond, M.; Kolios, A. Condition Monitoring Systems: A Systematic Literature Review on Machine-Learning Methods Improving Offshore-Wind Turbine Operational Management. *Int. J. Sustain. Energy* **2021**, *40*, 923–946. [CrossRef]
20. Zhou, X. 2D Vector Gravity Potential and Line Integrals for the Gravity Anomaly Caused by a 2D Mass of Depth-Dependent Density Contrast. *Geophysics* **2008**, *73*, I43–I50. [CrossRef]
21. De, S.; Dey, S.; Bhatia, S.; Bhattacharyya, S. Chapter 1—An Introduction to Data Mining in Social Networks. In *Advanced Data Mining Tools and Methods for Social Computing*; De, S., Dey, S., Bhattacharyya, S., Bhatia, S., Eds.; Hybrid Computational Intelligence for Pattern Analysis; Academic Press: Cambridge, MA, USA, 2022; pp. 1–25, ISBN 978-0-323-85708-6.
22. Han, C.; Yuan, Y.; Mei, T.; Geng, H. Data stream outlier detection algorithm based on K-means. *Comput. Eng. Appl.* **2017**, *3*, 58–63. (In Chinese)
23. Pan, H.; Xu, H.; Zheng, J.; Tong, J.; Cheng, J. Twin Robust Matrix Machine for Intelligent Fault Identification of Outlier Samples in Roller Bearing. *Knowl.-Based Syst.* **2022**, *252*, 109391. [CrossRef]
24. Degirmenci, A.; Karal, O. Efficient Density and Cluster Based Incremental Outlier Detection in Data Streams. *Inf. Sci.* **2022**, *607*, 901–920. [CrossRef]
25. Chen, L.; Wang, W.; Yang, Y. CELOF: Effective and Fast Memory Efficient Local Outlier Detection in High-Dimensional Data Streams. *Appl. Soft Comput.* **2021**, *102*, 107079. [CrossRef]
26. Li, K.; Gao, X.; Fu, S.; Diao, X.; Ye, P.; Xue, B.; Yu, J.; Huang, Z. Robust Outlier Detection Based on the Changing Rate of Directed Density Ratio. *Expert Syst. Appl.* **2022**, *207*, 117988. [CrossRef]
27. Zhao, Y.; Ye, L.; Zhu, Q. Characteristic and Processing Method of Abnormal Data Clusters Caused by Wind Curtailments in Wind Farms. *Autom. Electr. Power Syst.* **2014**, *38*, 40–46. (In Chinese)
28. Zhao, Y.; Ye, L.; Wang, W.; Sun, H.; Ju, Y.; Tang, Y. Data-Driven Correction Approach to Refine Power Curve of Wind Farm Under Wind Curtailment. *IEEE Trans. Sustain. Energy* **2018**, *9*, 95–105. [CrossRef]