Research Article

Study on Discrete Manufacturing Quality Control Technology Based on Big Data and Pattern Recognition

Chen Xin-chun, Ding Peng , Yan Nai-qing, and Bi Meng-xue

Jiangsu XCMG Construction Machinery Research Institute Ltd., Xuzhou 221004, China

Correspondence should be addressed to Ding Peng; 1123151292@qq.com

Received 3 September 2020; Revised 25 January 2021; Accepted 8 February 2021; Published 20 February 2021

Academic Editor: Adrian Neagu

Copyright © 2021 Chen Xin-chun et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Aiming at the quality control problems in the discrete manufacturing process of large and superlarge equipment, which cannot meet the urgent needs of production, a quality control method based on big data and pattern recognition is proposed. A large amount of data is collected through the test equipment developed in the discrete manufacturing process; a database of typical working conditions and an information tracking system relying on the cloud platform were formed. The working conditions were divided by the principal component analysis (PCA) and improved K-means algorithm. The Markov prediction model predicts the working conditions, recognizes the pattern with typical working conditions, regulates the processing parameters, and achieves quality control. Taking the quality control of the hydraulic cylinder manufacturing process above 5 m as an example for experimental verification, the experiments indicated that working conditions can be automatically identified and classified through pattern recognition technology. The process capability index Cpk increased from 0.6 to 1, which proved the effectiveness of quality control and the improvement of processing capabilities.

1. Introduction

Large and superlarge equipment is an indispensable product for important manufacturing industries such as construction machinery, aerospace, and steel. It is also an important foundation for the development of manufacturing and defense equipment. Complex systems, harsh environments, high loads, long-term operation, and relatively backward maintenance systems, equipment often encounter various failures, which seriously affect the progress and benefits of construction projects and the safety of people's property. Large and superlarge equipment is a typical discrete manufacturing process with multiple varieties and small batches. The production process is complex, the manufacturing is expensive, and the maintenance process is cumbersome. It is of great significance to improve the manufacturing quality control of large and superlarge equipment.

Now the order-driven production model has gradually transformed the manufacturing process of enterprises from traditional mass production to customized production of

multiple varieties and small batches. How to reduce costs and shorten the production cycle while meeting customer needs is a key issue. Manufacturing process optimization is the essence of manufacturing companies, and process management and quality control rely more on intelligent decision-making methods [1]. Scholars at home and abroad have conducted research on quality control during processing. Jiang XY [2] used similar manufacturing theory, statistical process control, neural network, and other theories and technologies to establish an intelligent process quality control system. Wang [3] proposed an intelligent quality control system modeling method based on multi-agent. Zheng [4] proposed a Bayesian network and big data analysis integration method for manufacturing process quality analysis and control. Wu [5] proposed a quality control method for complex product assembly process based on digital twin technology. Lee [6] developed a multivariate dynamic quality control theory and proposed an optimal design scheme based on Hotelling t^2 control chart and a control method with variable sampling interval. Celano [7] considered the limited sampling number of the small-batch

production processes and studied the monitoring performance and statistical properties of the t-control chart. Castagliola [8] studied the design method of the EWMA control chart in the case of unknown sample data, compared the different detection effects of the CUSUM control chart and the EWMA control chart, and summarized the applicable occasions of the two control charts. Jiang PY [9] proposed a multiprocess machining quality prediction modeling method based on an assignment error transfer network to perform machining quality control. Xiang C [10] analyzed several key technologies of data monitoring and quality control system based on the Internet of Things technology from the perspective of system integration. Hosseininasab S ME [11] used the clustering algorithm to mine and classify the quality data for quality control. At the same time, some scholars introduced data mining, brittleness theory, and other technologies into the process quality control [12, 13].

Due to the immature detection and quality control system of key quality control points in the manufacturing process, the quality control of large-scale discrete manufacturing cannot be applied in actual production. We take the actual production of hydraulic cylinders above 5 m as an example to carry out quality control of the processing process and establish the typical working conditions and pattern recognition in the actual processing process to carry out the whole process quality control. It is of great significance to improve processing accuracy and production efficiency and save processing time and production costs.

2. Quality Control Principles

With the development of information technology, big data has brought new opportunities for the development of all walks of life. As a typical discrete manufacturing enterprise, large and superlarge equipment has a large amount of data collected in the manufacturing process and a large amount of information. At the same time, there is often a high degree of correlation between variables. The data relationship presented is often nonlinear, and the subprocesses are coupled with each other. It makes the establishment of quality process control model more difficult. The lack of two technologies for typical working condition construction and complex processing condition recognition is the difficulty faced by the current large and superlarge equipment quality control technology. In view of such difficulties, this paper sets up various sensors, cameras, and other pieces of data collection equipment through the process, records all factors that affect the quality of processing, accumulates big data, and uses the improved *K-means* working condition algorithm to divide the working conditions and form typical working conditions. Control the quality of key quality control points through pattern recognition technology. In the production process, constantly measure the quality characteristics of the production system, and with the help of various quality statistical analysis methods and control methods, comparative analysis, quality decision-making and quality evaluation with the quality control objectives of key processes, and continuously feedback decision-making

information and evaluation results to various stages of the production system to perform quality control of the production process. The principle of quality control of product manufacturing process is shown in Figure 1.

3. Working Condition Pattern Recognition

3.1. Information Traceability. Network collaborative manufacturing is the only way for the transformation and upgrading of the manufacturing industry. It plays an important role in making full use of limited resources, improving the agility of manufacturing, and shortening the production cycle [14]. Make full use of advanced technologies such as cloud computing and big data analysis, with the help of various sensors, and achieve the acquisition of the entire processing data of large and superlarge equipment from the raw material plant-the external plant-equipment manufacturing plant through wireless transmission and network transmission technology to build a database. For the content of the general large and superlarge equipment database system, it should include subdatabases such as the intrinsic parameter database of the device, the knowledge base, the real-time dynamic database, and the historical database of the large and superlarge equipment. Integrate subdatabases into a whole database system for quality control. After the database is established, the corresponding settings of the database can be realized according to different application permissions, and the performance of application operations such as query and modification of the database can be repeatedly verified. Through the continuous adjustment and improvement of the database system, the data collected in the database will be further enriched and optimized to provide a solid guarantee for better playing the role of the database.

The data is shared on the cloud platform in the form of a database system. All processes within the company use a local area network-shared database to establish a company's internal quality control system. Various companies can rely on integrated information network, Beidou satellite communication, and so on and form a set of global connectivity security network system to achieve on-demand access in fixed places and mobile conditions by transmission encryption, isolation switching, and other technologies. Construct a cloud platform based on information integration of quality control system as shown in Figure 2.

In order to better control the quality and find key quality control points, manufacturing information traceability is required during the manufacturing process of large and superlarge equipment. Each product or material has a unique identification code. Each code corresponds to a quality control point, connecting the stations at the quality control point to form a production line. The collection of production lines of all products or materials constitutes a manufacturing network, which is a traceability model of processing and manufacturing process information. All information in the manufacturing process of large and superlarge equipment involved in the whole process can be traced through this model and the cloud platform architecture.

FIGURE 1: Quality control principle of the product manufacturing process.

FIGURE 2: Cloud platform architecture based on quality control system information integration.

3.2. PCA. Due to the large number of materials, processing, equipment, and other parameters designed during the manufacturing process of large and superlarge equipment, the correlation between the parameters is relatively complex, and it is not easy to predict and optimize the processing quality. In order to better perform quality control, this study discards some of the influencing factors according to the correlation between various parameters and the degree of impact on straightness by PCA [15] and further reduces the analysis index, so that the workload will be greatly reduced in the future.

According to the quality big data, establish the original parameter matrix composed of the data of materials, processing, and equipment (straightness is the number of rows, and the influence parameter is the number of columns).

$$R = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1n} \\ X_{21} & X_{22} & \dots & X_{2n} \\ \dots & \dots & \dots & \dots \\ X_{m1} & X_{n2} & \dots & X_{mn} \end{bmatrix}. \quad (1)$$

According to the PCA mathematics, the matrix $R = [X_{ij}]_{m \times n}$ is standardized, the covariance matrix is established, the eigenvalues are restored, the main elements of the matrix are sorted from left to right, the cumulative contribution rate is calculated, and the main elements are determined.

The original parameter matrix $R = [X_{ij}]_{m \times n}$ is standardized to obtain a matrix $X = [x^*_{ij}]_{m \times n}$, where

$$\begin{aligned} x^*_{ij} &= \frac{x_{ij} - \bar{x}_j}{s_j}, \quad i = 1, 2, \dots, m; j = 1, 2, \dots, n, \\ \bar{x}_j &= \frac{\sum_{i=1}^m x_{ij}}{m}, \\ s_j^2 &= \frac{\sum_{i=1}^m (x_{ij} - \bar{x}_j)^2}{m-1}. \end{aligned} \quad (2)$$

The correlation matrix C of the standardized matrix X is solved:

$$\begin{aligned} C &= (r_{ik})_{n \times n} \\ r_{jk} &= \frac{1}{m-1} \sum_{i=1}^m (z_{ij}, z_{ik}), \quad j, k = 1, 2, \dots, n. \end{aligned} \quad (3)$$

The eigenvalues of matrix C are solved and arranged from large to small as

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n. \quad (4)$$

The contribution rate of i -th principal component is calculated:

$$\rho_i = \frac{\lambda_i}{\sum_j \lambda_j}. \quad (5)$$

In general, the cumulative contribution rate exceeding 85% can meet the requirements.

3.3. Division of Operating Conditions Based on Improved K-Means Algorithm. According to the information tracing system established above, a large amount of large and superlarge equipment has accumulated data in the manufacturing process. The data presents a nonlinear relationship, which makes the data processing cumbersome. The improved *K-means* clustering algorithm is simple in principle, is easy to implement, and can quickly deal with a large amount of data. The clustering effect is better, and it is suitable for dividing the working conditions of the data.

The evaluation criterion function is defined as

$$G(k) = \sum_{i=1}^k \frac{\sum_{j=1}^{n_i} (x_{ij} - m_i)^2}{\sum_{l=1}^k \sum_{r=1}^{n_l} (x_{lr} - m_l)^2}. \quad (6)$$

In formula (6), x_{ij} is the j -th data of the i -th category; n_i is the number of the i -th data; m_i is the center of the i -th category; x_{lr} is the r -th data of the l -th category. In formula (6), the numerator is the distance between the data within the class and the center of the class, indicating the closeness of the class; the denominator is the sum of the distance between the

data outside the class and the center of the class, indicating the dispersion between different classes. The smaller the ratio, the more concentrated the data within the class, and the more dispersed the data among different classes, the more reasonable the division.

The specific steps of the *K-means* algorithm are as follows:

- (1) Select the initial value $k = n$
- (2) Cluster all the data to get all the clustering centers m_i ($i = 1, 2, \dots, k$) and the data x_{ij} ($i = 1, 2, \dots, k; j = 1, 2, \dots, n_i$)
- (3) Calculate $G_n(k)$ according to equation (6)
- (4) When $k = n+1$, repeat step (2); calculate $G_{n+1}(k)$ according to equation (6)
- (5) If $G_n(k) > G_{n+1}(k)$, repeat steps (3)-(4) to get $G_{n+2}(k)$; if $G_n(k) > G_{n+1}(k)$ iteration ends; otherwise, continue to repeat steps (3)-(5)

The selection of the initial clustering center m_i ($i = 1, 2, \dots, k$) has a great influence on the clustering result and running time, so it is necessary to select the appropriate k clustering centers. In view of this shortcoming, this paper optimizes the method of randomly initializing cluster centers, and the specific steps for improving the *K-means* algorithm to select cluster centers are as follows:

- (1) Randomly select data from the input data set as the first clustering center m_1
- (2) For each data x_{ij} in the data set, calculate the distance $D(x_{ij}) = \min \|x_{ij} - m_l\|^2$ from the nearest cluster center in the selected cluster center
- (3) Select the next new data as the new clustering center; the selection principle is the point with larger $D(x)$; the probability of being selected as the clustering center is greater
- (4) Repeat steps (3)-(4) until k cluster centers are selected

The improved K-means algorithm is used to divide the working conditions of key quality control point quality data in the discrete manufacturing process of large and superlarge equipment, and the processing parameters corresponding to the key quality control points are excavated to obtain typical processing conditions. By dividing the working conditions of the processing process, a database of typical processing conditions is constructed to lay the foundation for subsequent quality control.

4. Quality Control and Optimization

4.1. Markov Prediction Model. According to the big data during the processing quality of large and superlarge equipment, analyze the dynamic response characteristics of processing quality, and compare the prediction methods such as neural network and response surface method. Finally, it was decided to use Markov prediction theory to establish a key quality control point prediction model to improve the prediction accuracy of quality parameters and save calculation time [15].

Take each working condition divided in Part 3 as a state and record as $E_1, E_2, \dots, E_n$, a total of n possible states, and the working condition at time t is denoted as $x_t, x_t \in (E_1, E_2, \dots, E_n)$. Let P_{ij} be the state transition probability from state E_i to state E_j .

$$P = \begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1n} \\ P_{21} & P_{22} & \cdots & P_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ P_{n1} & P_{n2} & \cdots & P_{nn} \end{bmatrix}. \quad (7)$$

P is called the state transition probability matrix.

If the predicted straightness is currently in state E_i , then at the next moment, it may change from state E_i to any one of $E_1, E_2, \dots, E_i, \dots, E_n$. In order to calculate the state transition probability matrix P , the transition probability P_{ij} of each state transitioning to any other state is found. In order to find each P_{ij} , the idea of frequency approximation probability is used to calculate it.

$$P_{ij} = \frac{a_{ij}}{\sum_{j=1}^m a_{ij}}, \quad (8)$$

where a_{ij} is the number of observations from state E_i to state E_j .

The state probability $\pi_j(k)$ represents the probability that an event is in E_j at the k -th moment after k state transitions under the condition that the initial state ($k=0$) is known.

$$\sum_{j=1}^N \pi_j(k) = 1. \quad (9)$$

Markov model is shown as follows:

$$\pi(n) = \pi(0)P^n. \quad (10)$$

In the formula, $\pi(n)$ is the state probability vector at time n , and $\pi(0)$ is the initial state probability vector. According to the Markov model, the state of future straightness can be predicted.

4.2. Pattern Recognition and Optimization. As a typical discrete manufacturing product, the quality of large and superlarge equipment is mainly affected by raw material data, processing parameters, and equipment parameters during the manufacturing process. According to the quality data of key quality control points after processing, the working conditions are divided and the typical working condition database is established. Collect the data that affects the processing quality parameters in real time and compare it with the typical working condition database to verify whether it conforms to the existing historical typical working conditions of the working condition database. If it belongs to the current working condition, it prompts the category of the working condition; otherwise, save the current processing parameter data and automatically judge. The processing parameter data is automatically discriminated afterwards. If the processing quality results can meet

the processing requirements, a new typical working condition is established. Otherwise, the parameter adjustment is performed to conform to the typical working condition. A typical working condition database is constructed as shown in Figure 3.

$X=(x_1, x_2, \dots, x_n)$ represents an n -dimensional processing condition. And W typical processing conditions are $\omega_1, \omega_2, \dots, \omega_W$; the basic problem of pattern recognition is to find W discriminant functions $d_1(X), d_2(X), \dots, d_W(X)$ based on attributes. Working condition X belongs to typical processing working condition ω_i ; then

$$d_i(X) > d_j(X) \quad j = 1, 2, \dots, W; \quad j \neq i. \quad (11)$$

An unknown processing condition X is said to belong to the i -th typical processing condition, and it is only valid when X is substituted into all discriminant functions to obtain the maximum value of $d_i(X)$.

Define the prototype of a typical processing condition as the average value of the model.

$$m_j = \frac{1}{N_j} \sum_{X \in \omega_j} X_j, \quad j = 1, 2, \dots, W. \quad (12)$$

In the formula, W is the number of typical processing conditions, N_j is the number of parameters of a typical processing condition ω_j , and all these typical processing conditions are taken in the sum formula.

The method of determining the category member of an unknown processing condition X is to assign it to the nearest condition. The use of Euclidean spatial distance discrimination closeness simplifies the problem of calculating distance measures.

$$D_j(X) = \|X - m_j\|, \quad j = 1, 2, \dots, W. \quad (13)$$

When the value of $D_j(X)$ is the minimum distance, X is classified as a typical processing condition ω_i . The minimum distance indicated that it is the best match.

Choosing the minimum distance is equivalent to the following function evaluation:

$$d_j(X) = X^T m_j - \frac{1}{2} m_j^T m_j, \quad j = 1, 2, \dots, W, \quad (14)$$

And when d gets the maximum value, X is classified as class ω_i .

5. Experimental Verification

Taking the actual production process of hydraulic cylinders of more than 5 m as an example, the data in the information traceability system established by the entire processing process are divided into working conditions to form a typical working condition. At the same time, the entire process data of 30 hydraulic cylinders were tracked for quality control. And there was no obvious failure in the external ambient temperature and production equipment. Through the quality control methods proposed in this paper, quality control of key processes is carried out to verify the feasibility of the algorithm.

FIGURE 3: Typical working condition data workflow.

5.1. Division of Typical Working Conditions. The cylinder barrel and the piston rod are the key parts of the hydraulic cylinder and the key to the overall performance of the hydraulic cylinder. The spiral pattern and straightness are the key quality control points that affect the processing quality of the cylinder and piston. There are many parameters such as materials, processing, and equipment designed, the correlation between the parameters is relatively complex, and it is not easy to predict and optimize the processing quality during the manufacturing process. Therefore, the PCA in Section 3.2 is used for dimensionality reduction. At the same time, the improved *K-means* algorithm is used to divide the spiral pattern and straightness data in the database, and five cluster centers are selected. The results of the division of processing conditions were shown in Figure 4. According to the clustering results, we can get the range of quality control points for typical working conditions as shown in Table 1.

FIGURE 4: Division of typical working conditions.

5.2. Pattern Recognition and Optimization. According to the actual production of hydraulic cylinder processing, the straightness and spiral pattern are detected by using detection equipment such as sensors to accumulate a large amount of data. Markov can be used to predict the working conditions based on the above typical working conditions. If the working conditions can meet the quality requirements, the next process can be carried out normally; otherwise, the quality parameters should be optimized to reduce the speed of maintenance and scrapping.

Taking each typical working condition divided in Table 1 for each state, the state transition moment P can be obtained according to formula (8):

$$P = \begin{bmatrix} 0 & 0.5 & 0.5 & 0 & 0 \\ 0 & \frac{2}{3} & \frac{1}{3} & 0 & 0 \\ 0.25 & 0 & 0.25 & 0.5 & 0 \\ 0 & 0 & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \\ 0 & \frac{1}{5} & 0 & \frac{1}{5} & \frac{3}{5} \end{bmatrix}. \quad (15)$$

The measured straightness data are used as the initial state in the fourth working condition, so the initial state probability vector is

$$\pi(0) = [1 \ 0 \ 0 \ 0 \ 0]. \quad (16)$$

According to the Markov prediction model, the probability of occurrence of each working condition in the future 30 times is predicted. The first working condition is taken as an example as shown in Figure 5.

Markov prediction model is not related to the initial state. It can be found from Figure 5 that the probability of occurrence of the working condition at the 21st time is completely stabilized at a certain value. The probability of the 21st prediction is

$$\pi(21) = \pi(0)P^{21} = [0.0504 \ 0.1765 \ 0.2017 \ 0.4034 \ 0.1681]. \quad (17)$$

The Markov model can predict the probability of occurrence of typical working conditions, which provides a basis for process improvement and theoretical analysis of production and processing sites. In order to verify the possibility of quality prediction and pattern recognition and

TABLE 1: Division of typical working conditions.

Typical working conditions	Key quality control points	Quality range/mm	Solution
1	Spiral pattern	0.2–0.5	No optimization
	Straightness	0.4–0.9	
2	Spiral pattern	0.45–0.6	Parameter optimization of centerless grinding process
	Straightness	0.6–0.9	No optimization
3	Spiral pattern	0.45–0.7	Optimization of cold drawing and scraper process parameters
	Straightness	0.4–0.85	No optimization
4	Spiral pattern	0.3–0.65	Optimization of cold drawing and scraper process parameters
	Straightness	0.8–1.2	Straightening; centerless grinding process parameter optimization
5	Spiral pattern	0.2–0.5	No optimization
	Straightness	0.6–1.1	Parameter optimization of centerless grinding process

FIGURE 5: Division of typical working conditions.

TABLE 2: Tracking data of key quality control points.

Times	1	2	3	4	5	6	7	8	9	10
Spiral pattern (mm)	0.443	0.428	0.44	0.595	0.428	0.455	0.512	0.442	0.448	0.525
Straightness (mm)	0.63	0.8	0.38	0.86	0.48	0.42	0.53	0.38	0.52	0.62
Times	11	12	13	14	15	16	17	18	19	20
Spiral pattern (mm)	0.319	0.534	0.384	0.438	0.473	0.518	0.268	0.472	0.402	0.417
Straightness (mm)	0.44	0.77	0.89	0.39	0.76	0.75	0.77	0.61	0.93	1.08
Times	21	22	23	24	25	26	27	28	29	30
Spiral pattern (mm)	0.315	0.629	0.504	0.408	0.281	0.475	0.377	0.509	0.322	0.429
Straightness (mm)	0.31	0.49	1.08	0.75	1.09	1.32	1.08	0.8	0.85	0.76

optimization, track the entire production and processing process of 30 hydraulic cylinders and extract the two key quality control point data of spiral pattern and straightness, as shown in Table 2. Figure 6 shows the comparison of the predicted probability and actual probability of the working condition prediction model. It can be concluded that the prediction error rate of Markov's working condition prediction model is 9.40%, and the difference is small, which is within the acceptable range. The prediction error rate of the

model is 9.40%; the difference is small and within the acceptable range.

The Markov prediction method can obtain the probabilities that any working condition may appear in the future processing, and the working condition corresponding to the straightness and spiral pattern of the next moment can also be obtained by a one-step transfer matrix. Owing to the problem of straightness or spiral pattern out of tolerance of some workpieces, we have adjusted relevant parameters of

FIGURE 6: Comparison of working condition prediction.

FIGURE 7: Pattern recognition.

TABLE 3: Comparison of process capability analysis before and after quality control.

Key indicators	Before quality control			After quality control		
	Index requirements/mm	C_{pk}	Quality level	Index requirements/mm	C_{pk}	Quality level
Spiral pattern	0.3–0.5	0.57	3σ	0.2–0.4	1.03	4σ
Straightness	0.5–1.0	0.62	3σ	0.3–0.9	1.07	4σ

TABLE 4: Hydraulic cylinder fault feedback table.

Year	Number of leaks in hydraulic cylinder	Number of sounds of hydraulic cylinder	Total failures of hydraulic cylinder	Failure rate (%)
2018	47	6	1376	3.85
2019	21	1	915	2.4
2020	14	0	782	1.8

the influencing factors of straightness and spiral pattern in the process of hydraulic cylinder processing. According to the conclusion obtained from the PCA, only the parameters determined by the PCA need to be adjusted to achieve the purpose of regulating straightness.

The pattern recognition technology is used to recognize the working conditions after processing. The distribution of typical processing conditions of 30 sets of data was shown in Figure 7. According to the distribution results of typical

working conditions, the quality parameters of the processing procedures are optimized according to the solution in Table 1. This study uses the process capability analysis method to compare and analyze the quality control before and after quality control. Table 3 shows the results of process capability analysis before and after quality control.

Through comparison of process capability analysis results before and after quality control, the key index of the spiral pattern before quality control is 0.3–0.5 mm, the key

index of straightness is 0.5–1.0 mm, and the process capability index C_{pk} of processing equipment is about 0.6, which is currently at the level of 3σ . The key index of the spiral pattern after quality control is 0.2–0.4 mm, the key index of straightness is 0.3–0.9 mm, the process capability index of processing equipment is about 1 C_{pk} , which is at the level of 4σ , and the processing quality has been significantly improved.

In order to further verify the effectiveness of the quality control of the hydraulic cylinders, the hydraulic cylinders after quality control were tracked in the market, and the market fault feedback was counted. The feedback rate of hydraulic cylinder leakage and noise caused by deviation of straightness and spiral will be reduced from 3.85% in 2018 to 2.4% in 2019 and 2.4% in 2019 to 1.8% in 2020. As shown in Table 4, the market feedback rate continued to decline steadily.

6. Conclusions

This research focuses on the quality control of discrete manufacturing of large and superlarge equipment, carries out work such as the establishment of information traceability system, the construction of typical working conditions, pattern recognition, quality prediction, and parameter optimization, and draws the following conclusions:

- (1) Filter the processing parameters that affect the quality through the information tracing system and PCA method, and use the improved K-means algorithm to divide the quality data in the database to establish typical working conditions.
- (2) The probability of the occurrence of the working condition can be predicted through the Markov prediction model. The error rate is only 9.40%, which is within the acceptable range and meets the actual production and processing requirements by comparing the predicted result with the measured value.
- (3) The effectiveness of quality control has been verified by process capability analysis, of which the process capability index C_{pk} has been increased from about 0.6 to about 1, respectively. The effect is significant.
- (4) This research provides technical support for the optimization of discrete manufacturing quality control parameters and high-precision machining of large and superlarge equipment.

Notations

R :	The original parameter matrix
X_{ij} :	The parameters affecting processing quality
X :	The standardized matrix of matrix R
C :	The correlation matrix of matrix X
λ_i :	The eigenvalues of matrix C
ρ_i :	The contribution rate of i th principal component
$G(k)$:	The evaluation criterion function
x_{ij} :	The j th data of the i th category
n_i :	The number of the i th data
m_i :	The center of the i th category

x_{lr} :	The r th data of the l th category
$D(x_{ij})$:	The distance from x_{ij} to the nearest cluster center
E_i :	The state of working condition divided in Part 3
P_{ij} :	The state transition probability from state E_i to state E_j
P :	The state transition probability matrix
a_{ij} :	The number of observations from state E_i to state E_j
$\pi(n)$:	The state probability vector at time n
X :	The n -dimensional working condition
W :	The number of typical processing conditions
$d_w(X)$:	The discriminant functions based on attributes
m_j :	The average value of this processing model
N_j :	The number of parameters of a typical processing condition
$D_j(X)$:	The distance between the working condition and the average value of this type of mode.

Data Availability

Times 1 2 3 4 5 6 7 8 9 10 Spiral pattern/mm 0.443 0.428 0.44 0.595 0.428 0.455 0.512 0.442 0.448 0.525 Straightness/mm 0.63 0.8 0.38 0.86 0.48 0.42 0.53 0.38 0.52 0.62 Times 11 12 13 14 15 16 17 18 19 20 Spiral pattern/mm 0.319 0.534 0.384 0.438 0.473 0.518 0.268 0.472 0.402 0.417 Straightness/mm 0.44 0.77 0.89 0.39 0.76 0.75 0.77 0.61 0.93 1.08 Times 21 22 23 24 25 26 27 28 29 30 Spiral pattern/mm 0.315 0.629 0.504 0.408 0.281 0.475 0.377 0.509 0.322 0.429 Straightness/mm 0.31 0.49 1.08 0.75 1.09 1.32 1.08 0.8 0.85 0.76

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors gratefully acknowledge the financial support from the Natural Science Foundation of Jiangsu Province (BK20180175) and the Key R&D Program of Xuzhou of Jiangsu Province (KC18096).

References

- [1] H. Kagermann, J. Helbig, A. Hellinger et al., *Recommendations for Implementing the Strategic Initiative INDUSTRIE 4.0: securing the Future of German Manufacturing Industry; Final Report of the Industrie 4.0 Working Group*, Acatech, Berlin, Germany, 2013.
- [2] X. Jiang, S. J. Wang, K. Zhao, and W. S. Wang, "Intelligent process quality control system for networked manufacturing," *Journal of Mechanical Engineering*, vol. 46, no. 04, pp. 186–194, 2010.
- [3] Q. M. Wang, K. C. Liu, and H. Y. Gao, "Method of Modeling Intelligent Quality Control System Based on Multi-Agent Technology," *Beijing Institute of Technology*, vol. 32, no. 5, pp. 535–539, 2012.
- [4] M.-K. Zheng, X.-G. Ming, X.-Y. Zhang, and G.-M. Li, "MapReduce based parallel bayesian network for manufacturing quality control," *Chinese Journal of Mechanical Engineering*, vol. 30, no. 5, pp. 1216–1226, 2017.
- [5] Y. Wu, L. Y. Yao, and H. Xiong, "Quality control method of complex product assembly process based on digital twin

- technology,” *Computer Integrated Manufacturing Systems*, vol. 25, no. 6, pp. 1568–1575, 2019.
- [6] M. H. Lee and M. B. C. Khoo, “Design of a multivariate exponentially weighted moving average control chart with variable sampling intervals,” *Computational Stats*, vol. 29, no. 1-2, pp. 189–214, 2014.
 - [7] G. Celano, P. Castagliola, E. Trovato et al., “Shewhart and EWMA t control charts for short production runs,” *Quality and Reliability Engineering International*, vol. 27, no. 3, pp. 313–326, 2011.
 - [8] P. Castagliola and F. O. Figueiredo, “The median chart with estimated parameters,” *European Journal of Industrial Engineering*, vol. 7, no. 5, pp. 594–614, 2013.
 - [9] P. Y. Jiang, Y. Wang, H. F. Wang, and M. Zheng, “Quality prediction of multistage machining processes based on assigned error propagation network,” *Chinese Journal of Mechanical Engineering*, vol. 16, no. 10, pp. 160–170, 2010.
 - [10] C. Xiang and B. Li, “Research on ship intelligent manufacturing data monitoring and quality control system based on industrial Internet of Things,” *The International Journal of Advanced Manufacturing Technology*, vol. 107, no. 3-4, pp. 983–992, 2019.
 - [11] S. M. E. Hosseini nasab and M. J. Ershadi, “Optimization of the number of clusters: a case study on multivariate quality control results of segment installation,” *International Journal of Advanced Manufacturing Technology*, vol. 64, no. 5-8, pp. 1049–1055, 2013.
 - [12] G. Köksal, İ. Batmaz, and M. C. Testik, “A review of data mining applications for quality improvement in manufacturing industry,” *Expert Systems with Applications*, vol. 38, no. 10, pp. 13448–13467, 2011.
 - [13] Y. T. Qin, L. P. Zhao, and Y. Y. Yao, “Quality dynamic characteristics analysis based on brittleness theory in complex manufacturing processes,” *Computer Integrated Manufacturing Systems*, vol. 25, no. 6, pp. 2240–2249, 2019.
 - [14] L. Ren and M. L. Ren, “Manufacturing service composition method based on service weighted synergy network,” *Journal of Mechanical Engineering*, vol. 54, no. 16, pp. 70–78, 2018.
 - [15] X. C. Chen, P. Ding, C. Wang, and N. Q. Yan, “Study on straightness prediction of centerless grinding of slender rod based on PCA and Markov,” *Journal of Physics: Conference Series*, vol. 1633, Article ID 012157, 2020.