Article

Algorithm for Correlation Diagnosis in Multivariate Process Quality Based on the Optimal Typical Correlated Component Pair Group

Qing Niu *, Shujie Cheng and Zeyang Qiu

Department of Product Design, Lanzhou Jiaotong University, Lanzhou 730070, China; suniannan@sina.cn (S.C.); qiuzeyang@mail.lzjtu.cn (Z.Q.)
* Correspondence: liuqing@mail.lzjtu.cn

Abstract: Correlation diagnosis in multivariate process quality management is an important and challenging issue. In this paper, a new approach based on the optimal typical correlated component pair group (OTCCPG) is proposed. Firstly, the theorem of correlation decomposition is proved to decompose the correlation of all the quality components as serial correlations of component pairs, and then according to the transitivity of correlations of component pairs, the decomposition result is represented by a correlation set of typical correlated component pairs. Finally, an algorithm for OTCCPG based on the maximum correlation spanning tree (MCST) is proposed, and T^2 control charts to monitor the correlations of component pairs in OTCCPG are established to form the correlation diagnostic system. Theoretical analysis and practice prove that the proposed method could reduce the space complexity of the diagnostic system greatly.

Keywords: correlation diagnosis; correlation decomposition; quality component pairs; T^2 control chart

1. Introduction

With the development of the modern global market, the product's quality has been one of the key factors that greatly influence the competitiveness of enterprises. In the whole formation of the product's quality, process quality is the most basic aspect because the product's quality will be influenced by every process's quality, directly or indirectly, so process quality control is the essence of quality management in manufacturing.

The theory of statistical process control (SPC) and its tools, that is, the control chart proposed by Shewhart, achieved the goal of univariate process quality management [1–4]. However, in modern manufacturing, many processes contain more than one quality component. Due to the correlation of quality components, all components and their correlation need to be monitored simultaneously [5,6]. For example, in a bivariate process with quality $y = (y_1, y_2)^T$, 20 quality data points are collected to plot the mean control charts of y_1 and y_2, as shown in Figures 1 and 2. We can see that the variations of the two components are usually regular, i.e., with the increment of y_1, there is a monotonic increment of y_2. This regular variation of a monotonic increment within a component with the increment (or decrement) of another is defined as the correlation of y.

Although y_1 and y_2 are under control in Figures 1 and 2, this does not mean that y is in a stable state. At the 10th point, y_1 is decreasing while y_2 is increasing; this is opposite to the general variations of the two components, so it means that the correlation of y is abnormal. At the 20th point, although both y_1 and y_2 are decreasing, the decrement of y_2 is too heavy to exceed the normal range; this also means that the correlation of y is abnormal. The above inferences can be verified by plotting the T^2 control chart for y, as illustrated in Figure 3.
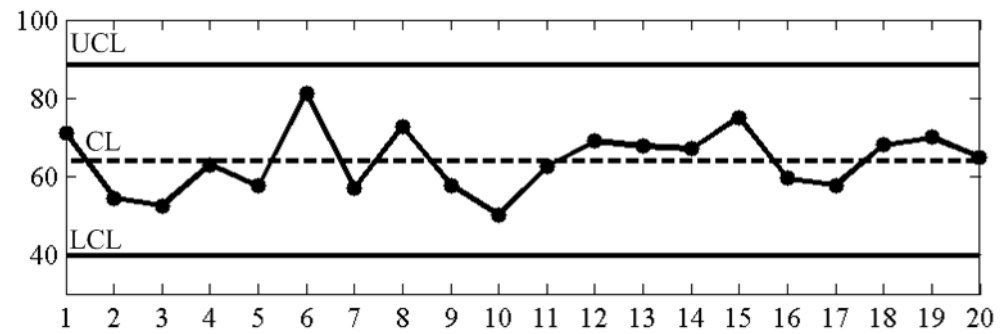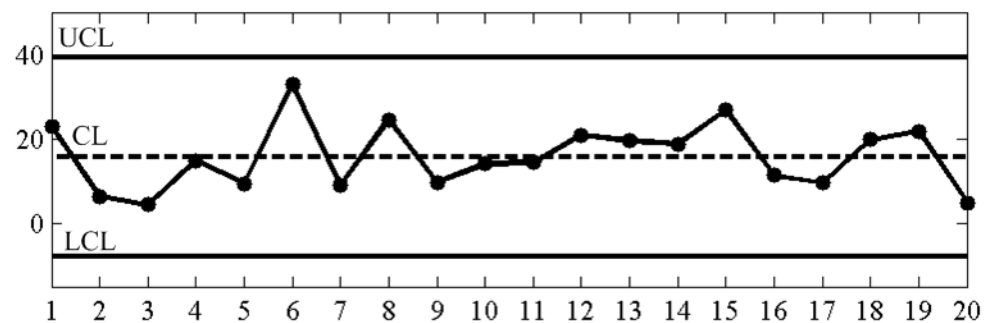
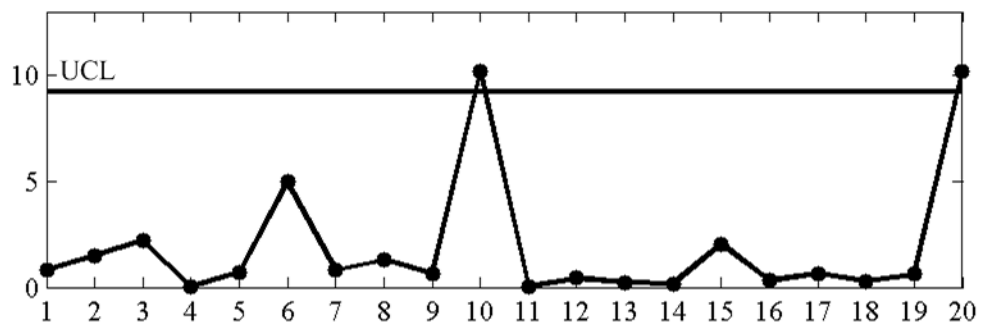**Figure 1.** Mean control chart of y_1.

Figure 2. Mean control chart of y_2.

Figure 3. T^2 control chart of $\boldsymbol{y}$.

The T^2 statistic was initially proposed by *Hotelling* to describe the correlation between quality components. For a p-dimensional process quality $\boldsymbol{y} = (y_1, y_2, \ldots, y_p)^T$, the T^2 statistic is defined as follows:

$$T^2 = (\boldsymbol{y} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{y} - \boldsymbol{\mu}) \tag{1}$$

where $\boldsymbol{\mu}$ is the mean vector, and $\boldsymbol{\Sigma}$ is the covariance matrix of $\boldsymbol{y}$. When T^2 > 0, it means that all the components in $\boldsymbol{y}$ are correlated.

It has been proven that the T^2 statistic follows the χ^2 distribution with p-freedom. Suppose α is the false probability, then the upper control limit (UCL) of the T^2 statistic is $\chi_\alpha^2(p)$, and the lower control limit (LCL) is 0. Thus, the T^2 control chart can be established to monitor the correlation shift of $\boldsymbol{y}$. However, this control chart cannot specify the cause(s) when the correlation shift is out of control.

1.1. Diagnosis Method Based on Component Combinations

In order to diagnose the cause(s) that cause the correlation shift to be out of control, an approach is to consider all the possible combinations of quality components when constructing their T^2 control charts. The cause(s) can then be specified by examining all the T^2 control charts when the correlation shift is out of control [7–9]. This approach is referred to as the component combination-based diagnosis (CCBD) method in this paper.

While this approach is theoretically sound and appealing, it has inherent deficiencies. For a p-dimensional process quality $\boldsymbol{y} = (y_1, y_2, \ldots, y_p)^T$, the number of T^2 control charts using the CCBD method is $N = C_p^2 + C_p^3 + \cdots + C_p^p = 2^p - p - 1$, where N is an exponential function of p and the space complexity is $O(2^p)$. When p increases, N will increase sharply, leading to a significant expansion of the diagnostic system scale, so this method is difficult to apply in practice.

On the other hand, the defect of information redundancy in diagnosis results cannot be avoided in the CCBD method. For example, in a four-dimensional process quality $\boldsymbol{y} = (y_1, y_2, y_3, y_4)^T$, suppose the abnormal correlation shift between y_1 and y_2 is the only cause that causes the correlation of $\boldsymbol{y}$ to be out of control. Now in the CCBD method, besides the T^2 control chart to monitor the correlation shift of (y_1, y_2) being out of control, the other combinations that contain y_1 and y_2, namely (y_1, y_2, y_3) and (y_1, y_2, y_4) are both out of control. This phenomenon is called the redundancy of diagnostic messages because the correlation shift of one component combination is out of control, and the correlations of other component combinations that contain the abnormal component combination are all out of control. The redundancy in diagnosis results is a disturbance for process quality adjustment.

1.2. Diagnosis Method Based on Principal Component Analysis

To reduce the scale of the correlation diagnostic system, scholars proposed that, by using the principal component analysis (PCA) method [10–13], the original process quality $\boldsymbol{y} = (y_1, y_2, \ldots, y_p)^T$ is converted to p-independent principal components and sorted by variance-decreasing order, denoted as $\boldsymbol{z} = (z_1, z_2, \ldots, z_p)^T$. Then, firstly, p Shewhart control charts are constructed to monitor the normality of z_i. Secondly, the first $n(n < p)$ principal components whose cumulative sum of their variance exceeds a specified critical value are grouped as component pairs, and T^2 control charts are constructed to monitor the normality of (z_i, z_j) $(i, j \leq n, i \neq j)$; At last, the normality of the rest of the principal component group $(z_{n+1}, z_{n+2}, \ldots, z_p)$ is monitored by a T^2 control chart.

The number of control charts based on the PAC method is $N = p + C_n^2 + 1 = p + n(n-1)/2 + 1$. In general, n increases while p increases, the space complexity is approximately $O(p^2)$, and the scale of the diagnostic system is still large. Furthermore, since z_i generally has no engineering significance after conversion, the cause(s) that cause the correlation shift of $\boldsymbol{y}$ out of control can only be specified by a comprehensive analysis of all the results in control charts, and by consulting the mapping relationship between $\boldsymbol{y}$ and $\boldsymbol{z}$, the calculation of the diagnosis is increased. Meanwhile, the redundancy of diagnostic messages also cannot be avoided.

1.3. Diagnosis Method Based on Correlation and Orthogonal Decomposition

In 1995, Mason, Young, and Tracy [14–16] proposed to decompose the T^2 statistic into conditional and unconditional terms by regression analysis, in which the conditional and unconditional terms have equal weight in the decomposition results and are orthogonally independent of each other. Then, according to the statistical distribution of the conditional and unconditional terms, the corresponding control limits are established, so as to diagnose the specific cause(s) when the manufacturing process is abnormal.

As an example, in bivariate process quality $\boldsymbol{y} = (y_1, y_2)^T$, the basic idea of the *MYT* orthogonal decomposition method [17–20] is to decompose the T^2 statistic into the following form:

$$T^2 = T_1^2 + T_{2\bullet1}^2 \tag{2}$$

where T_1^2, called the unconditional term, is related only to the quality component y_1 and is used to measure the contribution shift in y_1 to the T^2 statistic, and $T_{2\bullet1}^2$, called the conditional term, whose value is related to the conditional probability $P(y_2 \,|\, y_1)$, is used to measure the contribution in the correlation between y_1 and y_2 to the T^2 statistic.

Similar to Equation (2), the T^2 statistic can also be decomposed into another form as follows:

$$T^2 = T_2^2 + T_{1\bullet 2}^2 \tag{3}$$

where the unconditional term T_2^2 is related only to the quality component y_2 and is used to measure the contribution shift in y_2 to the T^2 statistic; the conditional term $T_{1\bullet 2}^2$ depends on the conditional probability $P(y_1 | y_2)$ and is used to measure the contribution in the correlation between y_2 and y_1 to the T^2 statistic.

Conditional probability $P(y_2 | y_1) \neq P(y_1 | y_2)$ under the condition that the quality components y_1 and y_2 are correlated, and hence the conditional term $T_{2\bullet 1}^2 \neq T_{1\bullet 2}^2$. Therefore, Equations (2) and (3) represent two different decomposition results of the T^2 statistic. In general, for a p-dimensional process quality $\boldsymbol{y} = (y_1, y_2, \ldots, y_p)^T$, the decomposition results have a total of $p!$. As the number of quality components increases, under the condition that every possible form of decomposition is analyzed, this will lead to a significant increase in calculations and a serious reduction in diagnostic efficiency. At the same time, the accuracy of the diagnostic results based on this method will be affected when there are obvious correlations between different quality components.

1.4. Intelligent Diagnosis Methods

In addition to the traditional diagnostic methods based on mathematical model analysis, in recent years, with the development of artificial intelligence technology, intelligent diagnostic methods have been applied to the field of multivariate process quality diagnosis, and the diagnostic methods based on artificial neural networks (ANNs) [21–23], Bayesian networks [24–26], support vector machines (SVMs) [27–29], etc. have been widely applied. Intelligent diagnostic methods can effectively reduce the scale of the diagnostic system and improve the diagnostic efficiency; however, these methods generally require a large amount of data to train the network's parameters, and the constructed network is generally suitable for specific applications, so their generality will be greatly restricted. Therefore, establishing a general and efficient method for multivariate process quality correlation diagnosis is a major problem to be solved in the field of quality management.

1.5. Sketch of the Algorithm

In this paper, a new correlation diagnosis method based on the optimal typical correlated component pair group is proposed. For the multivariate process quality $\boldsymbol{y} = (y_1, y_2, \ldots, y_p)^T$, the correlation decomposition theorem is first proved by drawing on the idea of decomposing the T^2 statistic in the *MYT* orthogonal decomposition method, which decomposes the correlation of all the quality components into the correlations of all the component pairs, to reduce the space complexity of the diagnostic system to $O(p^2)$. Then the monotonicity of the correlation between two quality components is used to prove the transitivity of the correlations of different component pairs, thus proving that the set formed by the correlations of all the component pairs is an equivalence relation. Next, drawing on the graphical representation of the equivalence relation in set theory and the concept of the minimum weighted spanning tree in graph theory, the computation method of the optimal correlated component pairs group is proposed, and reference is made to the diagnostic algorithm of component combination and the idea of combination of quality components in principal component analysis in order to represent the correlations of all the quality component pairs with the correlations in the optimal correlated component pairs group, so as to reduce the space complexity of the diagnostic system to $O(p)$.

2. Covariance Matrix Properties of Multivariate Process Quality

In the manufacturing process, factors affecting the product's quality can be attributed to five aspects: man, machines, materials, methods, and environment (4M1E). On this basis, ISO9000 supplemented another three factors: manufacturing software, auxiliary materials, and utilities. Among the many factors affecting the product's quality, changes in any one of them will have an impact on the final quality of the product, so the product's quality is

fluctuating in manufacturing. Tolerance theory is a direct proof of the fluctuation of the product's quality.

For the multivariate process quality $y = (y_1, y_2, \ldots, y_p)^T$, the covariance matrix is an important parameter to describe its correlation. Combined with the fluctuation of the product's quality in the manufacturing process, this paper first establishes three theorems describing the characteristics of the covariance matrix of multivariate process quality.

Theorem 1: *In the covariance matrix* $\boldsymbol{\Sigma}$ *of the multivariate process quality* $y = (y_1, y_2, \ldots, y_p)^T$, *all of the elements are not 0.*

Proof. Suppose the mean vector of y is $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_p)^T$. According to the definition of the covariance matrix, it is known that:

$$
\begin{aligned}
\boldsymbol{\Sigma} &= \begin{bmatrix}
E[(y_1 - \mu_1)(y_1 - \mu_1)] & E[(y_1 - \mu_1)(y_2 - \mu_2)] & \cdots & E[(y_1 - \mu_1)(y_p - \mu_p)] \\
E[(y_2 - \mu_2)(y_1 - \mu_1)] & E[(y_2 - \mu_2)(y_2 - \mu_2)] & \cdots & E[(y_2 - \mu_2)(y_p - \mu_p)] \\
\vdots & \vdots & \vdots & \vdots \\
E[(y_p - \mu_p)(y_1 - \mu_1)] & E[(y_p - \mu_p)(y_2 - \mu_2)] & \cdots & E[(y_p - \mu_p)(y_p - \mu_p)]
\end{bmatrix} \\
&= \begin{bmatrix}
\sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\
\sigma_{21} & \sigma_{21} & \cdots & \sigma_{2p} \\
\vdots & \vdots & \vdots & \vdots \\
\sigma_{p1} & \sigma_{p1} & \cdots & \sigma_{pp}
\end{bmatrix}
\end{aligned}
\tag{4}
$$

For any element $\sigma_{ij} = E[(y_i - \mu_i)(y_j - \mu_j)]$ in $\boldsymbol{\Sigma}$, the sufficient and necessary condition for it to be 0 is as follows:

$$
y_i = \mu_i \text{ or } y_j = \mu_j \tag{5}
$$

According to the properties of mathematical expectation, Equation (5) implies that the quality component y_i or y_j is a constant in the manufacturing process. Obviously, this contradicts the viewpoint of the fluctuation of the product's quality, and therefore, Equation (5) does not hold, i.e., all of the elements in $\boldsymbol{\Sigma}$ are not 0. □

Theorem 2: *The covariance matrix* $\boldsymbol{\Sigma}$ *of the multivariate process quality* $y = (y_1, y_2, \ldots, y_p)^T$ *is a real symmetric positive definite matrix.*

Proof. According to Equation (4) on the definition of the covariance matrix:

$$
\sigma_{ij} = E[(y_i - \mu_i)(y_j - \mu_j)] \tag{6}
$$

$$
\sigma_{ji} = E[(y_j - \mu_j)(y_i - \mu_i)] \tag{7}
$$

Compared to Equations (6) and (7), the properties of mathematical expectation can be seen as follows:

$$
\sigma_{ij} = \sigma_{ji} \tag{8}
$$

Equation (8) shows that $\boldsymbol{\Sigma}$ is a symmetric matrix.
Let p-dimensional vector $c = (c_1, c_2, \ldots, c_p)^T \neq \mathbf{0}$.

$$
\mathbf{c}^T \boldsymbol{\Sigma} c = (c_1, c_2, \cdots c_p) \boldsymbol{\Sigma} (c_1, c_2, \cdots c_p)^T \tag{9}
$$

Bringing Equation (4) into (9), after simplification and consolidation, we obtain the following:

$$
\mathbf{c}^T \boldsymbol{\Sigma} \mathbf{c} = E\left[\left(\sum_{i=1}^{p} c_i (y_i - \mu_i) \right) \left(\sum_{k=1}^{p} (y_k - \mu_k) c_k \right) \right] \tag{10}
$$

Let the random variable $z = \sum\limits_{i=1}^{p} c_i(y_i - \mu_i)$. Bringing this into Equation (10), we obtain the following:

$$c^T \boldsymbol{\Sigma} \mathbf{c} = E(z^2) \geq 0 \tag{11}$$

From the proof of Theorem 1, it is clear that, according to the viewpoint of the fluctuation of the product's quality, $z \neq 0$, i.e.,

$$\mathbf{c}^T \boldsymbol{\Sigma} c = E(z^2) > 0 \tag{12}$$

Equation (12) shows that the covariance matrix $\boldsymbol{\Sigma}$ of the multivariate process quality $\boldsymbol{y} = (y_1, y_2, \ldots, y_p)^T$ is a real symmetric positive definite matrix. $\square$

Theorem 3: *The inverse matrix $\boldsymbol{\Sigma}^{-1}$ of the covariance matrix $\boldsymbol{\Sigma}$ of the multivariate process quality $\boldsymbol{y} = (y_1, y_2, \ldots, y_p)^T$ is a real symmetric positive definite matrix.*

Proof. First, prove the symmetry of $\boldsymbol{\Sigma}^{-1}$. It follows from the symmetry of $\boldsymbol{\Sigma}$:

$$\boldsymbol{\Sigma} = \boldsymbol{\Sigma}^T \tag{13}$$

Inverting both ends of Equation (13), we obtain the following:

$$\boldsymbol{\Sigma}^{-1} = \left(\boldsymbol{\Sigma}^T\right)^{-1} = \left(\boldsymbol{\Sigma}^{-1}\right)^T \tag{14}$$

The above equation shows that $\boldsymbol{\Sigma}^{-1}$ is a symmetric matrix.

Let the eigenvalues of $\boldsymbol{\Sigma}$ be $\lambda_1, \lambda_2, \ldots, \lambda_p$. By the positive definiteness of $\boldsymbol{\Sigma}$, $\lambda_i > 0$ ($i \leq p$). According to the nature of the inverse matrix, the eigenvalues of $\boldsymbol{\Sigma}^{-1}$ are $1/\lambda_1, 1/\lambda_2, \ldots, 1/\lambda_p$, i.e., the eigenvalues of $\boldsymbol{\Sigma}^{-1}$ are all greater than 0, so $\boldsymbol{\Sigma}^{-1}$ is a positive definite matrix. $\square$

3. Correlation Diagnosis Method Based on OTCCPG

In the CCBD method, the exponential function between N and p is the main reason for the difficulty of applying this approach. If the gradient of N with p can be lowered by proper means, the defect of the diagnostic system scale expanding greatly while p is increasing will be avoided to a certain extent, and thus this approach can be applied in multivariate process quality management.

3.1. Correlation Decomposition

Theorem 4: *In the multivariate process quality $\boldsymbol{y} = (y_1, y_2, \ldots, y_p)^T$, the sufficient and necessary condition of the correlation of all components exists is that for any two components y_i and y_j, they are correlated.*

Proof. Firstly, the sufficiency of Theorem 4 is proved. Any two components y_i and y_j in $\boldsymbol{y}$ that are correlated show that $\sigma_{ij} \neq 0$. From Theorems 2 and 3, the covariance matrix $\boldsymbol{\Sigma}$ and its inverse matrix $\boldsymbol{\Sigma}^{-1}$ are real symmetric positive definite matrices. From the definition of the T^2 statistic in Equation (1), it is clear that for any sample data, its T^2 statistic is greater than 0, i.e., the correlation of all components exists.

The following proves the necessity of Theorem 4 by reductio ad absurdum. The existence of a correlation between all the components in $\boldsymbol{y}$ implies that for any sample data, its T^2 statistic is greater than 0. From the definition of the T^2 statistic in Equation (1), there exists an inverse matrix of the covariance matrix $\boldsymbol{\Sigma}$ of $\boldsymbol{y}$, and the rank of $\boldsymbol{\Sigma}$ is p.

$$R(\boldsymbol{\Sigma}) = p \tag{15}$$

Assume y_k and y_j in $\boldsymbol{y}$ are uncorrelated, i.e., $\sigma_{kj} = 0$. By the definition of covariance, there are, as follows:

$$\sigma_{kj} = E[(y_k - \mu_k)(y_j - \mu_j)] = 0 \tag{16}$$

The sufficient and necessary condition for Equation (16) to hold is $y_k = \mu_k$ or $y_j = \mu_j$. It may be useful to set $y_k = \mu_k$. From the definition of covariance, we know that for any component y_i ($i = 1, 2, \ldots, p$), there are, as follows:

$$\sigma_{ki} = E[(y_k - \mu_k)(y_i - \mu_i)] = 0 \tag{17}$$

Equation (17) shows that in the covariance matrix $\boldsymbol{\Sigma}$ of $\boldsymbol{y}$, the kth row and kth column are both 0, i.e., $R(\boldsymbol{\Sigma}) \leq p - 1$. This contradicts Equation (15), the assumption is not valid, and the necessity of Theorem 4 is proved. $\square$

Theorem 4 means that the correlation of all the quality components can be represented as correlations of component pairs, so in the correlation diagnostic system, it only needs to monitor the correlation shifts of all the component pairs. In addition, a T^2 control chart to monitor the correlation shift of all the components should be added. The number of T^2 control charts is $N = C_p^2 + 1 = p(p - 1)/2 + 1$, where N is the power function of p. The space complexity of the diagnostic system is lowered to $O(p^2)$. Compared to the CCBD method, the gradient of N with p has decreased significantly. Meanwhile, because the component pair is the minimum combination of components, information redundancy in diagnosis results can be avoided.

3.2. Transitivity of Component Pairs' Correlations

Although the functional relation between N and p is lowered to a power function by correlation decomposition, N will still increase rapidly while p is increasing, so further proper ways should be adopted to reduce the scale of the diagnostic system on the basis of the above analysis.

Definition 1: *In the multivariate process quality $\boldsymbol{y} = (y_1, y_2, \ldots, y_p)^T$, $Y = \{(y_i, y_j) \mid i, j = 1, 2, \ldots, p, i \neq j\}$ is the component pair set, and Y^* is another set that meets the following two properties*:

1. *$Y^* \subseteq Y$;*
2. *Any component in $\boldsymbol{y}$ should appear at least once in Y^*.*

Then Y^* is defined as a typical correlated component pair group (TCCPG) of $\boldsymbol{y}$.
Based on the above definition, there is Theorem 5 shown below.

Theorem 5: *In the multivariate process quality $\boldsymbol{y} = (y_1, y_2, \ldots, y_p)^T$, suppose Y^* is a TCCPG of $\boldsymbol{y}$, then the correlation of any component pair (y_i, y_j) ($i, j = 1, 2, \ldots, p, i \neq j$) can be represented by the correlations of the typical correlated component pairs in Y^*.*

Proof. The proof of Theorem 5 is given below in two cases.

(1) $(y_i, y_j) \in Y^*$.

It is obvious that Theorem 5 holds.

(2) $(y_i, y_j) \notin Y^*$.

The analysis of Figures 1–3 shows that for any two components y_j and y_k, the correlation of them means that with an increment in y_j, there is a proper and monotonically increasing increment in y_k; therefore, the correlation between y_j and y_k can be described by a function $y_k = f(y_j)$ that is monotonous in its domain. For any sample data, if the variations of y_j and y_k violate the monotonicity of this function or the magnitude of the change is beyond the reasonable range limited by this function, it means that the correlation between y_j and y_k is out of control.

When $(y_i, y_j) \notin Y^*$, from the second property in Definition 1, we know that there is a component pair chain (y_i, y_k), (y_k, y_m), $\ldots$, (y_s, y_j) in Y^*. Suppose the correlations of

these component pairs are denoted as $y_i = f_{ik}(y_k)$, $y_k = f_{km}(y_m)$, ..., $y_j = f_{sj}(y_j)$, then from the properties of compound function, the correlation between y_i and y_j can be written as $y_i = f_{ik}[f_{km}[...[f_{sj}(y_j)]...]$, i.e., the correlation between y_i and y_j can be represented by the correlations of component pairs in Y^*, Theorem 5 is hold. □

Theorem 5 means that in multivariate process quality, the correlations of component pairs have the property of transitivity, i.e., any correlation between two different components can be represented by the correlations of component pairs in TCCPG. Corresponding to the manufacturing process, if all the correlations of component pairs in TCCPG are normal, then the rest of the correlations of component pairs are normal, too; the correlation of all the quality components is under control. When the correlation of all the quality components is out of control, one or more correlations of component pairs in TCCPG must be abnormal. If these abnormal correlations are adjusted to normal levels by proper means, it can be assured that the correlation of all the quality components should be recovered to normal levels.

According to the above analysis, the correlation control of all the component pairs can be converted to monitor the ones in TCCPG. In addition, a control chart to monitor the correlation of all the quality components should be added to form the correlation diagnostic system.

3.3. Maximum Correlated Spanning Tree

In the above analysis, when using Theorem 5 to establish the diagnostic model of multivariate process quality correlations, the key problem is to solve the TCCPG. Here, the method of solving the TCCPG is given on the basis of the principles in set theory and graph theory.

Theorem 6: *In the multivariate process quality* $\boldsymbol{y} = (y_1, y_2, ..., y_p)^T$, *the set of binary correlations of all the quality component pairs is an equivalence relation.*

Proof. In set theory, a binary relation set with properties of self-reflexivity, symmetry, and transitivity is defined as an equivalence relation. Obviously, any component y_i in $\boldsymbol{y}$ and itself must be correlated, and the correlation coefficient is 1, so the self-reflexivity holds. If components y_i and y_j are correlated, then y_j and y_i are also correlated, and the quantitative description of their correlation, i.e., the correlation coefficients, are equal, i.e., $\rho_{ij} = \rho_{ji}$, so the symmetry property holds. According to Theorem 5, the binary correlations of all the component pairs have the property of transitivity. In summary, Theorem 6 holds. □

The set of equivalence relations can be characterized graphically [30–32]. For the multivariate process quality $\boldsymbol{y} = (y_1, y_2, ..., y_p)^T$, the set of vertices $V = \{y_1, y_2, ..., y_p\}$ is built with all the quality components. If y_i and y_j are correlated, an undirected edge e_{ij} is drawn to describe the binary correlation between y_i and y_j, and an undirected graph G representing the binary correlations between all the quality components can be obtained. From Theorem 1, any two quality components in $\boldsymbol{y}$ are correlated, and therefore, the undirected graph G contains $p(p-1)/2$ edges. G is a completely undirected graph.

In the completely undirected graph G, the correlation between any two components y_i and y_j can be described by the connectivity between the vertices y_i and y_j in V. According to the basic properties of completely undirected graphs, a loop L can be found in G, which contains all the vertices in V. If only one edge e_{ij} in L is deleted, the vertices in V are still connected, and the correlation between y_i and y_j can be represented by the correlations of the other vertices in L-e_{ij}. From our knowledge of graph theory [33,34], we know that L-e_{ij} is a spanning tree of G.

From Definition 1, the component pair group consisting of two vertices of each edge in the spanning tree of G is a TCCPG of $\boldsymbol{y}$. Therefore, the problem of solving the TCCPG of $\boldsymbol{y}$ can be converted into the problem of solving a spanning tree in the correlation graph G of $\boldsymbol{y}$.

The spanning tree of G is not unique in general. To find the most typical correlated component pairs group, take the correlation coefficient matrix $P = (\rho_{ij})_{p \times p}$ of y as a reference, set ρ_{ij} as the weight of the edge e_{ij} in G, and then a weighted correlation graph of y is obtained, which is denoted as G^*. Now the problem of finding the most typical correlated component pair group is converted to the one that calculates the maximum weighted spanning tree of G^*. This spanning tree is named the maximum-correlated spanning tree (MCST).

Considering that G^* is a dense graph with more edges, take the Prim algorithm [35–37] as the reference. The maximum correlated spanning tree of y can be calculated by the following steps:

(1) Let $U = \{y_1\}$, $V = \{y_1, y_2, \ldots, y_p\}$, and $E = \Phi$;
(2) Let e_{ij} denote the edge whose absolute value of the weight is maximized and its vertices $y_i \in U$, $y_j \in V - U$;
(3) Add y_j to U, e_{ij} to E;
(4) If $U \neq V$, go to step (2); otherwise, the algorithm is over.

When the algorithm is executed, the collection of vertex pairs of every edge in E is a TCCPG of y, and the binary correlations of all the quality component pairs can be represented to the maximum extent by this component pairs group, so it is called the optimal typical correlated component pairs group (OTCCPG).

3.4. Correlation Diagnosis Method

There are $p - 1$ edges in a spanning tree of a completely undirected graph with p vertices. Therefore, the number of component pairs in the OTCCPG method is $p - 1$. Establish T^2 control charts for every component pair in the OTCCPG to form the correlation diagnostic system; its space complexity is $O(p)$.

The above analysis is founded on the condition that the mean vector μ, covariance matrix Σ, and correlation coefficient matrix P of the manufacturing process are given. However, in many applications, these parameters are generally unknown. In this case, the unbiased estimator of the manufacturing process parameters can be calculated from a set of sample data S_i ($i = 1, 2, \ldots, n$) collected while the process is in a stable state.

$$\mu = \mu_1 \mu_2 \cdots \mu_p = \frac{1}{n} \sum_{i=1}^{n} S_i \tag{18}$$

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{21} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{p1} & \sigma_{p1} & \cdots & \sigma_{pp} \end{pmatrix} = \frac{1}{n-1} \sum_{i=1}^{n} (S_i - \mu)^T (S_i - \mu) \tag{19}$$

$$P = \begin{pmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{21} & \rho_{21} & \cdots & \rho_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ \rho_{p1} & \rho_{p1} & \cdots & \rho_{pp} \end{pmatrix} \tag{20}$$

$$\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii} \sigma_{jj}}} \quad (i, j = 1, 2, \cdots, p) \tag{21}$$

Combining the above analyses, for the multivariate process quality $y = (y_1, y_2, \ldots, y_p)^T$, the correlation diagnostic system can be formed by the following steps:

(1) Collect sufficient quality data S_i ($i = 1, 2, \ldots, n$) while the manufacturing process is in a stable state;
(2) Calculate the manufacturing parameters according to Equations (18)–(21);
(3) Take the coefficient matrix P as a reference and draw the weighted correlation graph G^* of y;

(4)   Find the MCST of G^*;

(5)   Obtain the OTCCPG of y according to the MCST;

(6)   For every component pair (y_i, y_j) in the OTCCPG, establish its T^2 control chart, K_{ij};

(7)   Establish the T^2 control chart K to monitor the correlation shift of all the quality components.

The number of T^2 control charts in this diagnostic system is p. When K is normal, the correlation of all the quality components is under control. When an out-of-control signal is generated in K, the root cause(s) can be specified by examining all the rest of the T^2 control charts.

4. A Case Study

The following case shows how the diagnostic system works in practice. The automotive purifier coating production line is mainly composed of three key steps: pre-treatment, coating, heating, and curing. Pre-treatment is the surface treatment of the parts to be coated to remove oil and rust; the coating process needs to ensure that the coating coverage is uniform and the thickness is consistent to avoid defects such as too thin, too thick, and leakage of the coating; and then curing of the coating material by heating. In the coating process, in order to ensure the uniformity of the coating thickness, the pressure, mixing temperature, coating temperature, slurry mass, and *pH* value should be controlled, which is denoted as $y = (y_1, y_2, y_3, y_4, y_5)^T$. Moreover, 20 quality data points are collected, as shown in Table 1.

Table 1. Collected sample data.

No.	Pressure y_1 (Pa)	Mixing Temperature y_2 (°C)	Coating Temperature y_3 (°C)	Slurry Mass y_4 (kg)	pH Value y_5
1	3940	40.7	40.7	9.98	2.26
2	3947	40.7	40.4	9.54	2.11
3	3950	41.2	40.7	9.09	2.04
4	3955	40.7	40.4	9.09	2.11
5	3958	41.2	40.7	9.98	2.26
6	4035	41.2	39.9	9.54	2.11
7	4038	41.7	40.1	8.19	1.97
8	4040	41.2	39.9	9.09	2.04
9	4043	41.7	40.1	9.09	2.19
10	4133	41.2	39.3	9.98	2.19
11	4137	41.7	39.6	9.98	2.33
12	4139	41.2	39.3	8.19	1.9
13	4221	42.2	39.1	8.64	1.97
14	4223	42.2	39.1	9.54	2.11
15	4226	42.2	39.1	9.54	2.19
16	4229	41.7	39.1	9.09	2.19
17	4319	42.7	38.5	9.98	2.26
18	4412	43.8	38	8.64	2.04
19	4417	43.3	38.3	8.19	1.97
20	4505	43.3	37.7	8.64	1.97

In order to analyze the reliability of the sample data, for each quality component, the probability of false alarm is taken as $\alpha = 0.0027$ according to the 3σ principle of the Shewhart control chart, and then the probability of false alarm in the correlation shift is taken as $\alpha_y = 0.025$ with reference to Bonferroni's inequality to make Shewhart control charts for the five quality components and T^2 control charts to monitor the correlation shift of the five quality components, as shown in Figures 4–9.
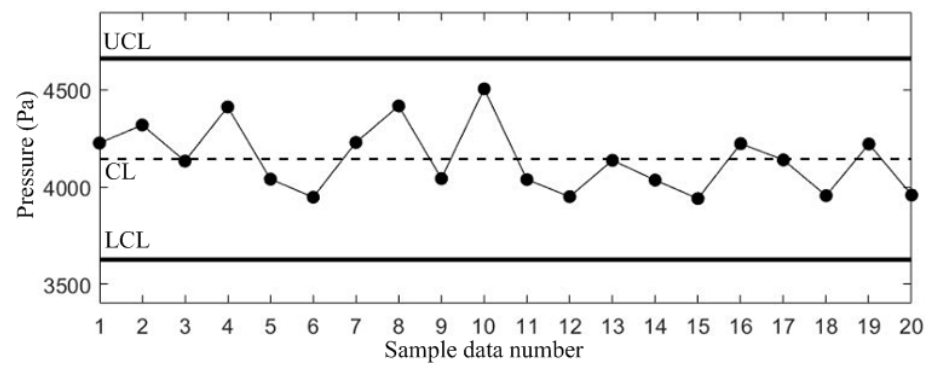
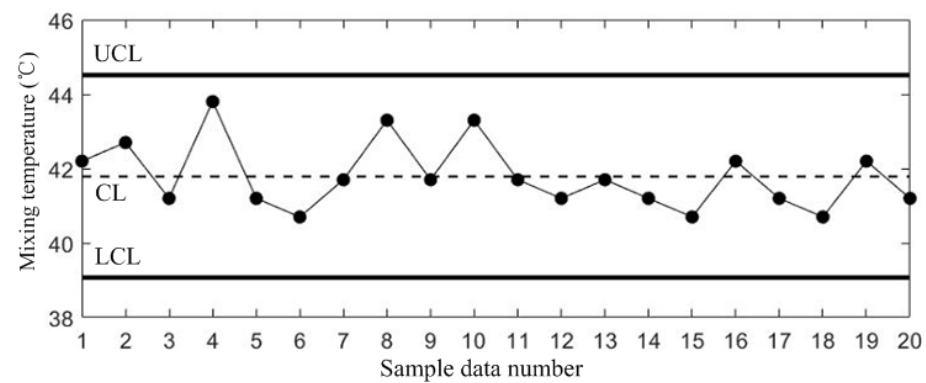**Figure 4.** Shewhart control chart of y_1 to monitor the mean shift of y_1 in Table 1.

Figure 5. Shewhart control chart of y_2. to monitor the mean shift of y_2 in Table 1.
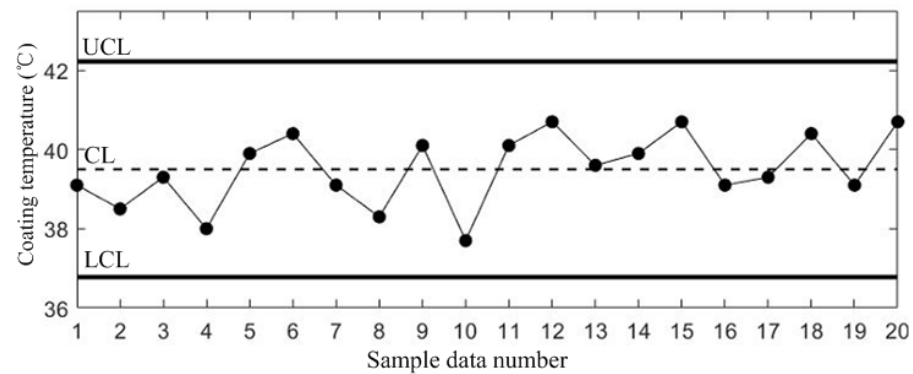

Figure 6. Shewhart control chart of y, to monitor the mean shift of y_3 in Table 1.
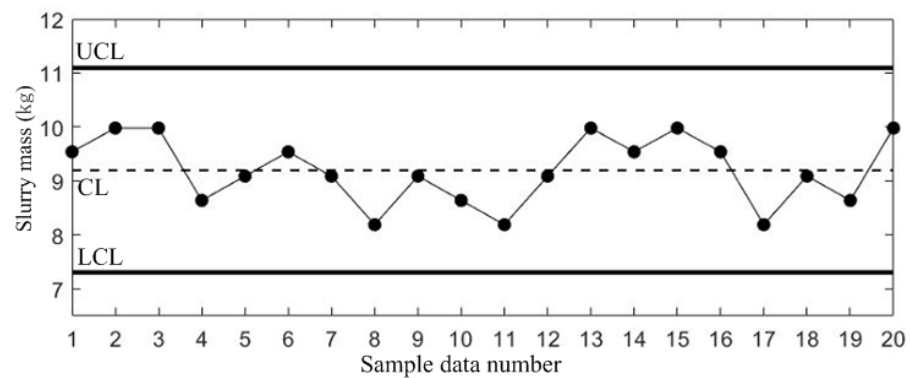

Figure 7. Shewhart control chart of y_4 to monitor the mean shift of y_4 in Table 1.
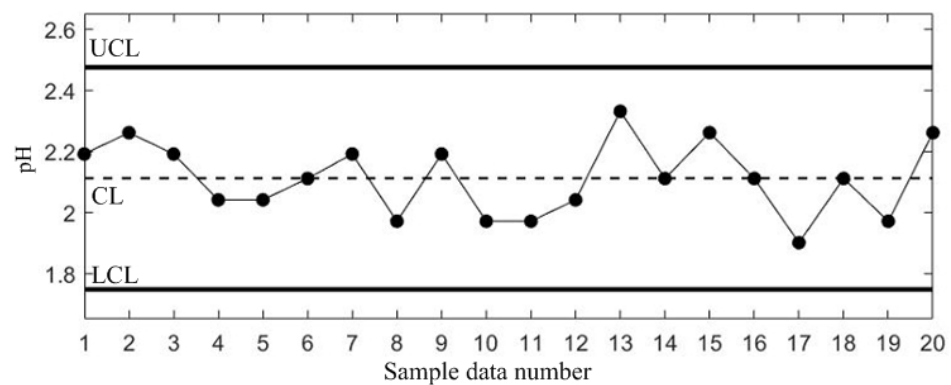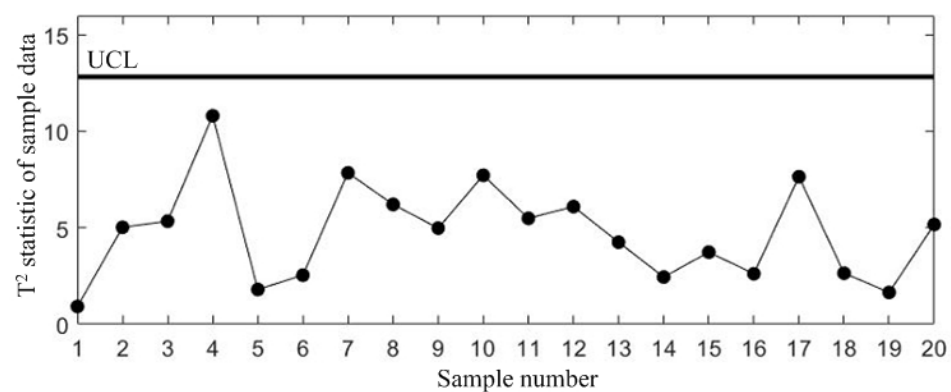
Figure 8. Shewhart control chart of y_5 to monitor the mean shift of y_5 in Table 1.

Figure 9. T^2 control chart of y to monitor the correlation shift of $(y_1, y_2, y_3, y_4, y_5)$ in Table 1.

Figures 4–9 show that the process is in a stable state; therefore, the data in Table 1 can be used to estimate the parameters of the coating process using Equations (18)–(21):

$$\mu = (4143,\ 41.8,\ 39.5,\ 9.2,\ 2.11)$$

$$\Sigma = \begin{pmatrix} 29795.6079 & 145.6247 & -154.6684 & -38.6679 & -5.8386 \\ 145.6247 & 0.8220 & -0.7342 & -0.2207 & -0.02986 \\ -154.6684 & -0.7342 & 0.8232 & 0.1983 & 0.0330 \\ -38.6679 & -0.2207 & 0.1983 & 0.3982 & 0.0685 \\ -5.8386 & -0.0298 & 0.0330 & 0.0685 & 0.0147 \end{pmatrix}$$

$$P = \begin{pmatrix} 1 & 0.9305 & -0.9876 & -0.3550 & -0.2793 \\ 0.9305 & 1 & -0.8926 & -0.3857 & -0.2712 \\ -0.9876 & -0.8926 & 1 & 0.3463 & 0.3003 \\ -0.3550 & -0.3857 & 0.3463 & 1 & 0.8959 \\ -0.2793 & -0.2712 & 0.3003 & 0.90 & 1 \end{pmatrix}$$

Take the coefficient matrix P as a reference; the weighted correlation graph G^* is shown in Figure 10, and then the MCST of G^* can be found, as shown in Figure 11.

According to Figure 11, the OTCCPG of y is $Y^* = \{(y_1, y_2), (y_1, y_3), (y_2, y_4), (y_4, y_5)\}$. For every component pair in Y^*, suppose the false probability $\alpha = 0.025$. K_{12} denotes the T^2 control chart to monitor the correlation shift between y_1 and y_2. K_{13} denotes the T^2 control chart to monitor the correlation shift between y_1 and y_3. Similar symbols are applied to denote the rest of the T^2 control charts. Finally, T^2 control chart K to monitor the correlation shift of all the quality components is established. These five T^2 control charts constitute the correlation diagnostic system for y.
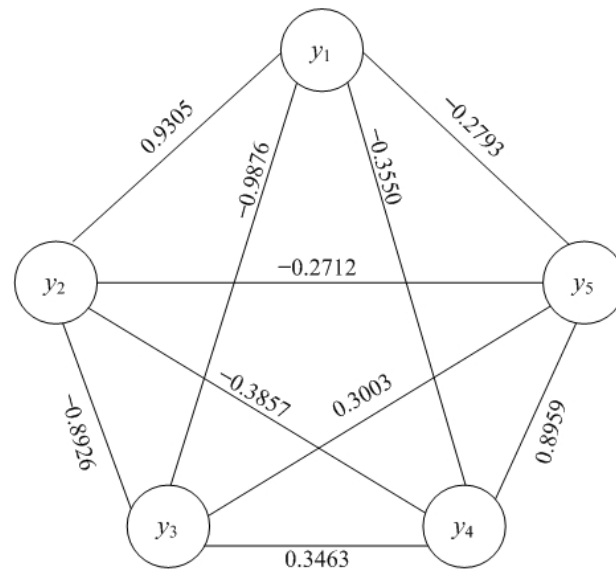
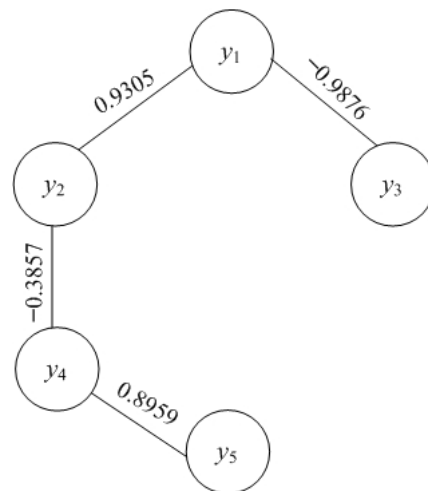**Figure 10.** Weighted correlation graph.

Figure 11. Maximum correlation spanning tree.

Now this diagnostic system can be used to monitor and diagnose the correlation shift of quality components. In subsequent manufacturing, four quality data points at different moments are collected, as shown in Table 2. The T^2 control chart of these four quality data points is drawn in K, as shown in Figure 12. We can see that at the last three points, the correlation shift of all the components is out of control.

Table 2. Test data collected in subsequent manufacturing.

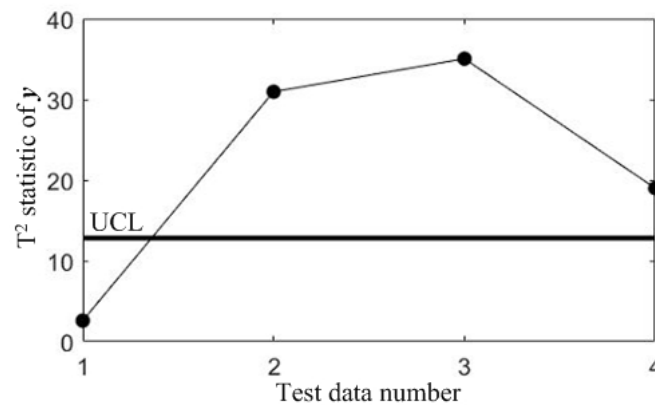| No.	Pressure y_1 (Pa)	Mixing Temperature y_2 (°C)	Coating Temperature y_3 (°C)	Slurry Mass y_4 (kg)	pH Value y_5
1	4226	42.2	39.1	9.98	2.26
2	4133	40.2	39.1	8.64	2.04
3	3947	40.7	39.9	8.19	1.97
4	4226	42.2	39.1	9.54	1.97

Figure 12. T^2 control chart K to monitor the correlation shift of $(y_1, y_2, y_3, y_4, y_5)$ in Table 2.

In order to diagnose the root cause(s) that cause the last three points to be out of control, check the rest of the T^2 control chart $(K_{12}, K_{13}, K_{24}, K_{45})$, as shown in Figures 13–16. From these charts, it can be concluded that the root causes that cause the second point to be out of control are the correlations of (y_1, y_2) and (y_1, y_3) being abnormal. The abnormal correlation of (y_1, y_3) causes the third point out of control, and the abnormal correlation of (y_4, y_5) leads to the last point out of control.
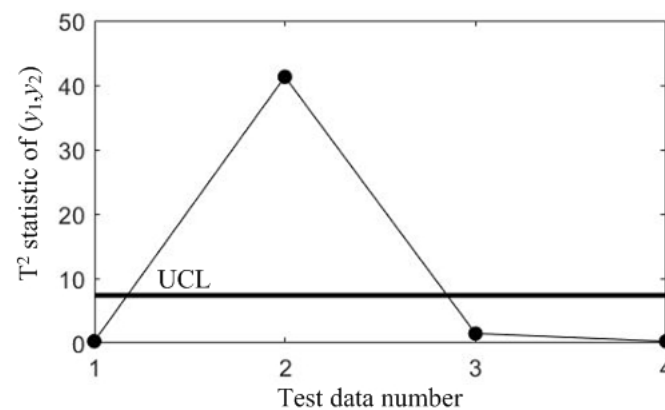
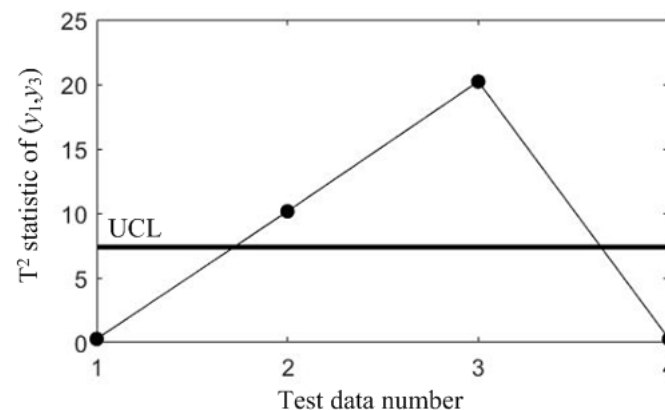**Figure 13.** T^2 control chart K_{12} to monitor the correlation shift of (y_1, y_2) in Table 2.

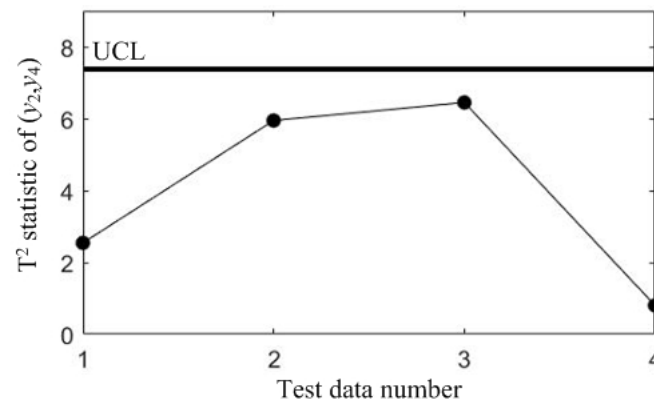**Figure 14.** T^2 control chart K_{13} to monitor the correlation shift of (y_1, y_3) in Table 2.

Figure 15. T^2 control chart K_{24} to monitor the correlation shift of (y_2, y_4) in Table 2.
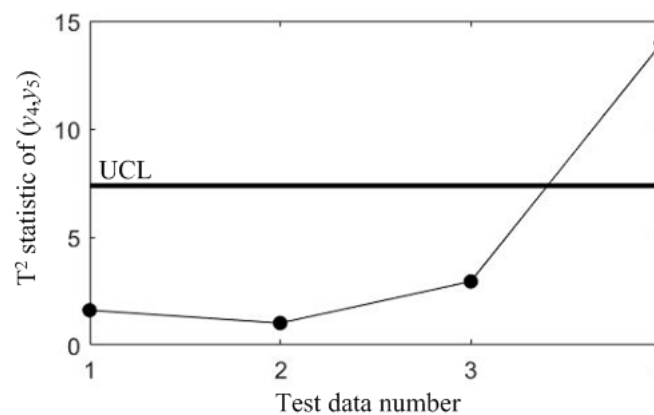

Figure 16. T^2 control chart K_{45} to monitor the correlation shift of (y_4, y_5) in Table 2.

In order to validate the above conclusions, 26 T^2 control charts contained in the diagnosis system using the CCBD method are constructed. The diagnostic results of the two systems are shown in Table 3.

Table 3. Comparison of the two diagnosis results.

No.	Diagnostic System Based on OTCCPG	Diagnostic System Using CCBD Method
1	Normal	Normal
2	(y_1, y_2), (y_1, y_3) are abnormal	(y_1, y_2), (y_1, y_3), (y_1, y_2, y_3), (y_1, y_2, y_4), (y_1, y_2, y_5), (y_1, y_3, y_4), (y_1, y_3, y_5), (y_1, y_2, y_3, y_4), (y_1, y_2, y_3, y_5), (y_1, y_2, y_4, y_5), (y_1, y_3, y_4, y_5) are abnormal
3	(y_1, y_3) is abnormal	(y_1, y_3), (y_1, y_2, y_3), (y_1, y_3, y_4), (y_1, y_3, y_5), (y_1, y_2, y_3, y_4), (y_1, y_2, y_3, y_5), (y_1, y_3, y_4, y_5) are abnormal
4	(y_4, y_5) is abnormal	(y_4, y_5), (y_1, y_4, y_5), (y_2, y_4, y_5), (y_3, y_4, y_5), (y_1, y_2, y_4, y_5), (y_1, y_3, y_4, y_5), (y_2, y_3, y_4, y_5) are abnormal

In Table 3, the conclusions are different except at the first point; this difference can be explained properly. At the second point, because the correlations of (y_1, y_2) and (y_1, y_3) are abnormal, the other correlations of component combinations that contain (y_1, y_2) or (y_1, y_3) must be abnormal, so they are redundant diagnostic results. If the redundant diagnostic results are removed, the conclusions of the two diagnostic systems are completely consistent at this point. A similar analysis could explain the difference at the other points, as shown in Table 4.

From the above comparison, we can see that the correlation diagnosis method of multivariate process quality based on OTCCPG can not only achieve the goal of specifying the cause(s) that cause the correlation shift out of control but could also avoid the redundancy

of diagnosis results, so it can provide more accurate diagnosis messages for the process quality adjustment in further steps.

Table 4. Comparison of the two diagnostic systems after redundant diagnostic results were removed.

No.	Diagnostic System Based on OTCCPG	Diagnostic System Using CCBD Method
1	Normal	Normal
2	(y_1, y_2), (y_1, y_3) are abnormal	(y_1, y_2), (y_1, y_3) are abnormal
3	(y_1, y_3) is abnormal	(y_1, y_3) is abnormal
4	(y_4, y_5) is abnormal	(y_4, y_5) is abnormal

5. Conclusions and Discussion

This paper proposes an OTCCPG-based correlation diagnosis method for multivariate process quality management. The development of this method included several steps. First, based on correlation decomposition, the correlation of all the quality components is decomposed as the ones of all the quality component pairs, and then, according to the transitivity of correlations of component pairs, the correlation of any component pair can be represented by the ones in TCCPG. Finally, the algorithm of the OTCCPG based on MCST is proposed, and the correlation diagnostic system based on the OTCCPG is illustrated. Compared with the existing diagnostic methods, the method proposed in this paper has the following advantages:

(1) The diagnosis is more efficient.

The space complexity of the multivariate process quality correlation diagnostic method based on OTCCPG is $O(p)$, while the space complexity of the diagnostic algorithm based on the CCBD method, the principal component analysis method, and the orthogonal decomposition of the T^2 statistic are $O(2^p)$, $O(p^2)$, and $O(p!)$, respectively. Therefore, the proposed method in this paper has higher diagnostic efficiency.

(2) The diagnostic results are more accurate.

The OTCCPG-based multivariate process quality correlation diagnosis method takes the correlation of component pairs as the diagnostic unit. Since component pairs are the smallest combination of quality components, the disadvantage of redundant diagnostic information in diagnostic algorithms based on the CCBD method, the principal component analysis method, and the orthogonal decomposition of T^2 statistics can be avoided to provide more accurate diagnostic results for manufacturing processes.

(3) Better generality.

Compared with the diagnostic methods based on artificial intelligence technology, the diagnostic method proposed in this paper is based on strict mathematical analysis as the theoretical foundation, which avoids the defect of intelligent diagnostic methods in which the network structure and parameters are oriented to specific applications. Therefore, the proposed method can be used as a general theoretical model for multivariate process quality correlation diagnosis.

The multivariate process quality correlation diagnostic method proposed in this paper is based on the fact that the manufacturing process parameters are known or can be estimated more accurately by using sufficient sample data. For the manufacturing process of large-batch products, it is feasible to make more accurate estimates of the manufacturing process parameters due to the sufficient available sample data for reference. However, for flexible manufacturing systems with variable varieties and small batches, the method proposed in this paper cannot be directly applied due to the small number of available samples. Considering the similarity of different batches of products in a flexible manufacturing system in terms of structure and process, the change of process parameters in flexible manufacturing should have certain regularity. Therefore, studying the changing laws of process parameters in the flexible manufacturing system and adapting the multivariate

process quality correlation diagnosis method based on mathematical statistical theory to the flexible manufacturing system can be further research in the field of quality management.

Author Contributions: Conceptualization, Q.N.; methodology, Q.N. and S.C.; resources, Z.Q.; writing—original draft, Q.N.; writing—review and editing, Q.N. and S.C.; supervision, Z.Q.; project administration, Q.N. and Z.Q. All authors have read and agreed to the published version of the manuscript.

References

1. Fernandes, F.H.; Ho, L.L.; Bourguignon, M. About Shewhart control charts to monitor the Weibull mean. *Qual. Reliab. Eng. Int.* **2019**, *35*, 2343–2357. [CrossRef]
2. Linda, L.H.; Fidel, H.F.; Roberto, C.Q. Improving Shewhart control chart performance for monitoring the Weibull mean. *Qual. Reliab. Eng. Int.* **2021**, *37*, 984–996.
3. Nguyen, H.D.; Tran, K.P.; Celano, G.; Maravelakis, P.E.; Castagliola, P. On the effect of the measurement error on Shewhart and EWMA control charts. *Int. J. Adv. Manuf. Technol.* **2020**, *107*, 4317–4332. [CrossRef]
4. Malela-Majika, J.C.; Shongwe, S.C.; Castagliola, P.; Mutambayi, R.M. A novel single composite Shewhart-EWMA control chart for monitoring the process mean. *Qual. Reliab. Eng. Int.* **2022**, *38*, 1760–1789. [CrossRef]
5. Mjimer, I.; Aoula, E.; Achouyab, E.H. Monitoring of overall equipment effectiveness by multivariate statistical process control. *Int. J. Lean Six Sigma* **2022**, *13*, 847–862. [CrossRef]
6. Sun, Y.; Younis, I.; Zhang, Y.; Zhou, H. Optimizing the quality control of multivariate processes under an improved Mahalanobis-Taguchi system. *Qual. Eng.* **2023**, *35*, 413–429. [CrossRef]
7. Bahrami, H.; Niaki, S.T.A.; Khedmati, M. Monitoring multivariate profiles in multistage processes. *Commun. Stat. Simul. Comput.* **2019**, *50*, 3436–3464. [CrossRef]
8. Muhammad, I.; Hong, L.D.; Fatima, S.Z. Incorporating principal component analysis into Hotelling T^2 control chart for compositional data monitoring. *Comput. Ind. Eng.* **2023**, *186*, 40–56.
9. Ershadi, M.J.; Niaki, S.T.A.; Azizi, A.; Esfahani, A.A.; Abadi, R.E. Monitoring data quality using Hoteling multivariate control chart. *Commun. Stat. Simul. Comput.* **2023**, *52*, 1591–1606. [CrossRef]
10. Huang, J.; Yan, X. Quality-driven principal component analysis combined with kernel least squares for multivariate statistical process monitoring. *IEEE Trans. Control Syst. Technol.* **2019**, *27*, 2688–2695. [CrossRef]
11. Qi, L.; Yi, X.; Yao, L.; Fang, Y.; Ren, Y. Quality-related process monitoring based on improved kernel principal component regression. *IEEE Access* **2021**, *9*, 132733–132745. [CrossRef]
12. Riaz, M.; Zaman, B.; Mehmood, R.; Abbas, N.; Abujiya, M. Advanced multivariate cumulative sum control charts based on principal component method with application. *Qual. Reliab. Eng. Int.* **2021**, *37*, 2760–2789. [CrossRef]
13. Sun, C.; Hou, J. An improved principal component regression for quality-related process monitoring of industrial control systems. *IEEE Access* **2017**, *5*, 21723–21730. [CrossRef]
14. Mason, R.L.; Tracy, N.D.; Young, J.C. Decomposition of T^2 for multivariate control chart interpretation. *J. Qual. Technol.* **1995**, *27*, 109–119. [CrossRef]
15. Mason, R.L.; Tracy, N.D.; Young, J.C. A practical approach for interpreting multivariate T^2 control chart signals. *J. Qual. Technol.* **1997**, *29*, 396–406. [CrossRef]
16. Mason, R.L.; Tracy, N.D.; Young, J.C. Improving the sensitivity of the T^2 statistic in multivariate process control. *J. Qual. Technol.* **1999**, *31*, 155–165. [CrossRef]
17. Akeem, A.A.; Yahaya, A.; Asiribo, O. Hotelling's T^2 decomposition: Approach for five process characteristics in a multivariate statistical process control. *Am. J. Theor. Appl. Stat.* **2015**, *4*, 432–437. [CrossRef]
18. Huang, X.H.; Xu, J.K.; Zhou, Q. Multi-scale diagnosis of spatial point interaction via decomposition of the k function-based T^2 statistic. *J. Qual. Technol.* **2017**, *49*, 213–227. [CrossRef]
19. Li, X.L.; Liu, S.S. Fault separation and detection algorithm based on Mason Young Tracy decomposition and Gaussian mixture models. *Int. J. Intell. Comput.* **2020**, *13*, 81–101. [CrossRef]
20. Ueda, R.M.; Souza, A.M. An effective approach to detect the source(s) of out-of-control signals in productive processes by vector error correction (VEC) residual and Hotelling's T^2 decomposition techniques. *Expert. Syst. Appl.* **2022**, *187*, 115979. [CrossRef]
21. Yu, J.B.; Zhang, C.Y.; Wang, S.J. Sparse one-dimensional convolutional neural network-based feature learning for fault detection and diagnosis in multivariable manufacturing processes. *Neural. Comput. Appl.* **2022**, *34*, 4343–4366. [CrossRef]
22. Samira, Z.; Moosa, A. Simultaneous fault diagnosis of wind turbine using multichannel convolutional neural networks. *ISA Trans.* **2021**, *108*, 230–239.

23. Xu, Q.Q.; Dong, J.; Peng, K.X. A novel method of neural network model predictive control integrated process monitoring and applications to hot rolling process. *Expert. Syst. Appl.* **2023**, *237*, 121682. [CrossRef]

24. Xian, X.C.; Li, J.; Liu, K.B. Causation-based monitoring and diagnosis for multivariate categorical processes with ordinal information. *IEEE Trans. Autom. Sci. Eng.* **2019**, *16*, 886–897. [CrossRef]

25. Rezki, N.; Kazar, O.; Mouss, L.H. A hybrid approach for complex industrial process monitoring. *J. Sci. Ind. Res. India* **2017**, *76*, 608–613.

26. Wang, Y.Z.; Liu, Y.; Khan, F. Semiparametric PCA and Bayesian network based process fault diagnosis technique. *Can. J. Chem. Eng.* **2017**, *95*, 1800–1816. [CrossRef]

27. Liang, J.P.; Zhang, K. A new hybrid fault diagnosis method for wind energy converters. *Electronics* **2023**, *12*, 1263. [CrossRef]

28. Zhang, H.Q.; Wang, J.C.; Wang, M. Integration of cuckoo search and fuzzy support vector machine for intelligent diagnosis of production process quality. *J. Ind. Manag. Optim.* **2022**, *18*, 195–217. [CrossRef]

29. Tang, J.; Zhao, Q.N. Motor rolling bearing fault diagnosis based on MVMD energy entropy and GWO-SVM. *J. Vibroeng.* **2023**, *25*, 1096–1107. [CrossRef]

30. Crone, L.; Fishman, L.; Jackson, S.C. Equivalence relations and determinacy. *J. Math. Log.* **2022**, *22*, 2250003. [CrossRef]

31. Zhuchok, Y.; Toichkina, O. Endotypes of partial equivalence relations. *Semigroup Forum* **2021**, *103*, 966–975. [CrossRef]

32. Frey, J. Categories of partial equivalence relations as localizations. *J. Pure Appl. Algebra* **2022**, *227*, 107115. [CrossRef]

33. Chakraborty, M.; Chowdhury, S.; Pal, R.K. Two algorithms for computing all spanning trees of a simple, undirected, and connected graph: Once assuming a complete graph. *IEEE Access* **2018**, *6*, 56290–56300. [CrossRef]

34. Alexander, V.S. Spanning tree of a multiple graph. *J. Comb. Optim.* **2022**, *43*, 850–869.

35. Zeng, X.X.; Liu, Q.L.; Yao, S. An improved Prim algorithm for connection scheme of last train in urban mass transit network. *Symmetry* **2019**, *11*, 681. [CrossRef]

36. Lukaszewski, A.; Nogal, L. Multi-sourced power system restoration strategy based on modified Prim's algorithm. *Bull. Pol. Acad. Sci. Tech. Sci.* **2021**, *69*, e137942. [CrossRef]

37. Łukaszewski, A.; Nogal, L.; Januszewski, M. The application of the modified Prim's algorithm to restore the power system using renewable energy sources. *Symmetry* **2022**, *14*, 1012. [CrossRef]