


---



---


## Profile Control Chart Based On Maximum Entropy

---


---

Authors: SEYEDEH AZADEH FALLAH MORTEZANEJAD   
– School of Automotive and Traffic Engineering, Jiangsu University,  
Zhenjiang, China  
azadeh.fallah@ujs.edu.cn

RUOCHEN WANG    
– School of Automotive and Traffic Engineering, Jiangsu University,  
Zhenjiang, China  
wrc@ujs.edu.cn

GHOLAMREZA MOHTASHAMI BORZADARAN   
– Ferdowsi University of Mashhad, department of Statistics,  
Mashhad, Iran  
grmohtashami@um.ac.ir

RENKAI DING   
– School of Automotive and Traffic Engineering, Jiangsu University,  
Zhenjiang, China  
drk@ujs.edu.cn

KIM PHUC TRAN   
– Univ. Lille, ENSAIT, ULR 2461, GEMTEX, Genie et Materiaux Textiles, F59000  
Lille, France;  
International Chair in DS & XAI, International Research Institute for Artificial  
Intelligence and Data Science, Dong A University,  
Danang, Vietnam  
kim-phuc.tran@ensait.fr

Received: Month 0000

Revised: Month 0000

Accepted: Month 0000

### Abstract:

- Monitoring a process over time is essential in manufacturing. In some cases, process quality is defined by various profile types. The goal is to monitor profile coefficients to detect minor, undesired shifts in the response variable. Traditional Control Charts (CCs) often fail to identify these small shifts. This paper addresses this issue by introducing a nonparametric CC based on the Maximum Entropy ( $ME$ ) principle for Linear Profile (LP) data. The novelty of this method lies in its use of  $ME$  to approximate the joint distribution of profile coefficients. A comprehensive simulation and two real data examples are presented to demonstrate the chart's applicability and efficient performance based on Average Run Length ( $ARL$ ).

### Keywords:

- *Linear Profile (LP); Maximum Entropy (ME); Control Chart (CC); Nonparametric Chart; Average Run Length (ARL).*

AMS Subject Classification:

- 49A05, 78B26.

---

## 1. Introduction

---

CCs are widely used for scrutinizing industry and manufacturing processes. Numerous articles were published on monitoring and detecting shifts in processes or systems, which are essential for effectively supervising equipment performance. Xie et al. (2000) published a paper on controlling process reliability, highlighting the strengths and weaknesses of various CCs, including Shewhart charts.

When various attributes of a process are monitored simultaneously, one of the most commonly used methods is the  $T^2$ -Hotelling CCs. For example, Wibawati et al. (2022) proposed a new version of the  $T^2$ -Hotelling CC based on a fuzzy neutrosophic concept to address the challenges associated with the traditional Hotelling charts, which require well-documented datasets. Imran et al. (2023) incorporated principal component analysis and the  $T^2$ -Hotelling CC to develop a robust technique for monitoring compositional data transformed by isometric log-ratio. Zaidi et al. (2023) combined the  $T^2$ -Hotelling statistic with a neural network to detect undesirable shifts, such as outliers and trends, for a classification task.

Profiles are used in situations where the qualities of products or processes are defined by a functional relationship between a response (dependent) variable and one or more explanatory (independent) variables. Regression models emphasize prediction and quantifying relationships, whereas profile models concentrate on describing and understanding the characteristics of specific groups. Roshanbin et al. (2022) instructed a multi-objective economic-statistical design for LPs. They considered three objective functions to model the cost: the Lorenzen-Vance cost function, the In-Control (IC)  $ARL_0$ , and the Out-Of-Control (OOC)  $ARL_1$ . Also, they established lower and upper bounds for  $ARL_0$  and  $ARL_1$ , respectively. Yeganeh and Shongwe (2023) conducted a study on profile monitoring of the cryptoclastic markets using a novel cryptocurrency charting application. They applied their proposed chart to monitor price variations for two of the most well-known cryptocurrencies, Bitcoin and Ethereum. To validate their chart, they performed parameter adjustments and compared the results with established technical indicators using a real dataset, ultimately demonstrating the effectiveness of their novel method. Nancy et al. (2023) conducted a survey on regression CC, which combines the conventional CC concept with regression methods for processes influenced by interrelated independent variables. Their paper facilitates the selection of the most suitable CC for researchers, based on the quality characteristics of the processes involved in their research needs.

Nonparametric charts are exceptional because they do not require any assumptions regarding the distribution of the variable. Consequently, misattributing an incorrect distribution of the data can lead to significant economic losses. As a result, extensive research has been conducted on the unknown distribution of data to identify changes in processes. Perdikis et al. (2023) examined three nonparametric Shewhart CCs for assessing scale parameters in finite-horizon production processes using nonparametric tests. A numerical analysis, along with a real-world industrial example, was provided to demonstrate the effectiveness and practical application of these techniques.

$ME$  possesses numerous outstanding features, including unbiasedness, suitability for ill-posed data, and effectiveness with small sample sizes, which render its results acceptable

across various applications. Additionally, *ME* has specific parameters that are learned from the dataset. [Dalleiger and Vreeken \(2020\)](#) introduced a new version of *ME* called relaxed *ME*, which utilizes dynamic factorization of the joint distribution to mitigate computational complexity. They emphasized several important applications of *ME*, including classification and pattern mining. [Che et al. \(2022\)](#) introduced a novel method for the decision-making trial and evaluation laboratory based on *ME* techniques to optimize the normalization of the direct matrix. This approach minimizes the information loss and enhances accuracy, making it suitable for emergency management. [Macedo \(2024\)](#) introduced a two-stage *ME* method for estimating model parameters and bootstrapping data replication in the time series regression. This approach showed significant advantages over traditional methods, as evidenced by both simulated and empirical examples. Here, we present another application of *ME* for detecting minor undesired shifts in manufacturing processes.

In this paper, we propose a novel monitoring method for production processes that moves beyond observing the corresponding mean. We utilize a straightforward LP approach for this new CC. The LP establishes a functional relationship between a response variable and a single explanatory variable, akin to a Linear Regression (*LR*) model. The objective of this paper is to determine an appropriate control limit for the profile coefficients. We introduce an innovative method for defining a CC based on a distribution-free framework. Therefore, instead of relying on the default normal distribution, we derive the unknown joint distribution of our model variables based on *ME*. The new CC is impressive when compared to the charts presented in this article for monitoring manufactured products. A crucial step in *ME* is to determine the unknown joint distribution of the available variables based on their observations. In real-world applications, the joint distribution is often passive and must be estimated effectively. Subsequently, we utilize this density to define an appropriate nonparametric CC. Finally, the unknown profile coefficients are approximated using this specified distribution. Besides, we apply the standard *LR* parameter estimators to compare with the newly proposed method. We incorporate the algorithms of *ME* and *LR* CCs to enhance the applicability and reproducibility of the results for users, along with a straightforward example demonstrating their application. Finally, we conduct simulation studies to compare the effectiveness of the explained CCs with the Fisher-based chart using the *ARL* as a metric. Ultimately, we apply these charts to real data examples to evaluate their performance in practice.

This paper has numerous applications across diverse challenges. For instance, [Geng et al. \(2023\)](#) developed a non-*LR* model to monitor the freshness of chicken meat. Similarly, [Zhao et al. \(2020\)](#) employed discrete element methods and fuzzy algorithms to analyze and control soil compaction during seedbed preparation. These are some examples of where our method can be effectively applied to facilitate decision-making.

The structure of this paper is organized as follows: In Section 2, the *ME* principle is employed to determine the unknown distribution functions. Section 3 presents relevant literature and fundamental definitions of profiles. In Section 4, both *ME* and *LR* are utilized to estimate the coefficients necessary for establishing control limits based on the  $T^2$ -Hotelling statistic. Simulation studies are discussed in Section 5. Section 6 provides an example using authentic data from the semiconductor production process. Section 7 features a real-world pharmaceutical example that demonstrates the effectiveness of *ME*. Finally, conclusions are drawn in Section 8.

---

## 2. Bivariate *ME* Distribution

---

Two significant principles in dynamic optimization are the Maximum Pontryagin (*MP*) principle, presented in [Kopp \(1962\)](#), and the Lagrange method. In this context, we utilize the *ME* principle, as introduced in [Jaynes \(1957\)](#), because our objective function is the Shannon entropy but the objective function in the Lagrange method can vary depending on the specific problem at hand. Although the *MP* and *ME* methods are related, they are used in different contexts and have distinct formulations tailored to their specific applications. The *MP* is specifically used in optimal control theory, where the goal is to optimize a performance criterion over time while considering the dynamics of a system. It incorporates the Hamiltonian, which combines the objective functional and the system dynamics. In contrast, the *ME* introduces Lagrange multipliers to incorporate constraints directly into the optimization process, allowing for the derivation of density functions that maximize entropy subject to the constraints. The *ME* is more general and can be applied to various types of optimization problems, particularly those involving static scenarios, whereas the *MP* is specifically designed for dynamic control systems where the evolution of states over time is critical. Thus, in many studies, including our case study, the results of the *ME* and *MP* can be seen as complementary, but they are not necessarily equivalent. Each method provides valuable insights within its own framework, depending on the nature of the problem being addressed.

In this paper we focus on Shannon entropy, introduced by [Shannon \(1948\)](#). Shannon entropy is widely utilized as an information criterion for approximating density functions. For instance, the distributions of financial variables, stock returns, and incomes can be derived from the *ME* by defining specific intentional constraints. The *ME* distribution is the most unbiased density that satisfies the given constraints. In statistical terms, an estimator is deemed unbiased if its mathematical expectation matches the true parameter it estimates. For a given distribution, this implies that when we calculate the expected value based on that distribution, it should equal the parameter the distribution is intended to estimate. This means that the *ME* distribution refers to the concept that the resulting distribution imposes the fewest additional assumptions beyond those specified in the constraints.

To determine the bivariate *ME*, let  $X$  and  $Y$  be two random variables with joint density function  $f_{X,Y}(x, y)$ . The joint Shannon entropy is defined as follows:

$$H(f) = - \int \int_{\mathbb{S}(X,Y)} \log(f_{X,Y}(x, y)) f_{X,Y}(x, y) dx dy,$$

where  $\mathbb{S}(X, Y)$  is the joint support set of  $X$  and  $Y$ .

Using *ME* to determine an unknown density function involves selecting constraints that accurately reflect our understanding of the process while remaining as non-informative as possible. The goal is to derive the distribution that best represents our state of knowledge without introducing additional bias. Understanding the modeling process and the available information are essential factors to consider when selecting appropriate constraints. This includes experimental data, known parameters such as the mean and variance, as well as any given physical laws. Knowledge of expected values, such as the mean or variance, can serve as constraints. If real data exists, it facilitates selection, and constraints can be established based

on the main characteristics associated with the data. Additionally, if further information is available, such as skewness or kurtosis, it can also be incorporated into the constraints. Since we investigate the *ME* for the approximation of density functions, one crucial constraint is normalization, which is defined as the integral of the function being equal to 1. Note that while adding more constraints may enhance the model, it also complicates the solution.

As a consequence, we establish constraints for  $X$  and  $Y$  based on mathematical expectations and the available data. Suppose  $m_i(\cdot)$ 's, for  $i = 1, \dots, r$ , are some arbitrary moment functions based on the available information from the data at hand. For example, one choice of  $m_i(x, y)$  can be  $\overline{xy} = n^{-1} \sum_{j=1}^n x_j y_j$ , where the value is computed using the existing dataset, and  $n$  represents the sample size. The functions  $h_i(X, Y)$ , for  $i = 1, \dots, r$ , are defined according to  $m_i(\cdot)$ 's. For instance, if  $m_i(x, y) = \overline{xy}$ , then its corresponding  $h_i(X, Y)$  is  $XY$ . So, *ME* constraints are defined as:

$$(2.1) \quad \begin{aligned} E(h_i(X, Y)|f) &= \int \int_{\mathbb{S}(X, Y)} h_i(X, Y) f_{X, Y}(x, y) dx dy \\ &= m_i(x, y), \quad i = 1, \dots, r, \end{aligned}$$

where  $r$  is the number of constraints and  $m_i(x, y)$ s are known values. In Equation (2.1),  $x$  and  $y$  represent the observations of  $X$  and  $Y$ , respectively, while  $f$  denotes their actual joint density function, which remains undetermined. Therefore, our objective is to approximate the function using the *ME*. To achieve this, we apply the Lagrange function, which is constructed by Shannon entropy and the constraints outlined in (2.1):

$$\begin{aligned} L(f, \lambda_0, \dots, \lambda_r) &= - \int \int_{\mathbb{S}(X, Y)} \log(f_{X, Y}(x, y)) f_{X, Y}(x, y) dx dy \\ &\quad - \lambda_0 \left\{ \int \int_{\mathbb{S}(X, Y)} f_{X, Y}(x, y) dx dy - 1 \right\} \\ &\quad - \sum_{i=1}^r \lambda_i \left\{ \int \int_{\mathbb{S}(X, Y)} h_i(x, y) f_{X, Y}(x, y) dx dy - m_i(x, y) \right\}. \end{aligned}$$

The coefficient  $\lambda_0$  ensures that the resulting function is a valid density function based on the following fact:

$$(2.2) \quad \int \int_{\mathbb{S}(X, Y)} f_{X, Y}(x, y) dx dy = 1.$$

So, by differentiation of the Lagrange function for  $f$ , we get:

$$\frac{\partial L(f, \lambda_0, \dots, \lambda_r)}{\partial f} = -\log(f) - 1 - \lambda_0 - \sum_{i=1}^r \lambda_i h_i(X, Y).$$

Then, by setting the equation equal to zero, the following equation must be solved for  $f$ :

$$\frac{\partial L(f, \lambda_0, \dots, \lambda_r)}{\partial f} = 0.$$

Finally, the *ME* density is derived:

$$(2.3) \quad f_{X, Y}(x, y) = \exp(-1 - \lambda_0 - \sum_{i=1}^r \lambda_i h_i(x, y)), \quad (x, y) \in \mathbb{S}(X, Y).$$

The Lagrange coefficients  $\lambda_0, \dots, \lambda_r$  are unknown and must be computed by substituting the function in Equation (2.3) into the desired constraints (2.1) and (2.2). Consequently, the *ME* coefficients are determined by solving the system of equations for the  $\lambda$  values. We provide a more detailed explanation in Subsection 5.1. In the following section, we aim to calculate the profile coefficients based on *ME* and compare them with those from *LR*.

---

### 3. Profile Definition

---

Profiles can take various forms, including simple LPs, multivariate LPs, multiple LPs, polynomial LPs, generalized LPs, and non-LPs. In this paper, we focus on simple LPs. First, let us assume that there are  $n$  fixed-point observations from variable  $X$ , and  $k$  random samples from variable  $Y$  corresponding to each observation of  $X$ . So, we have a vector of observations from  $X$  and a matrix of samples from  $Y$ . In an IC situation,  $Y$  and  $X$  are modeled with:

$$(3.1) \quad \underset{\sim}{y}_j = a + b \underset{\sim}{x} + \underset{\sim}{\varepsilon}_j, \quad j = 1, \dots, k,$$

where  $k$  represents the total number of samples, each of size  $n$ . The term  $\underset{\sim}{x}$  denotes the observations of  $X$ , and the intrinsic error  $\underset{\sim}{\varepsilon}_j$  is an independent random variable that follows a normal distribution with a mean of zero and a fixed variance of  $\sigma^2$ . The intercept  $a$  and slope  $b$  are referred to as profile coefficients. The model described in equation (3.1) is analogous to simple  $LR$ ; however, the key difference lies in the witnessing of vector  $X$ . In  $LR$ , there are different vectors of  $\underset{\sim}{y}_j$  and  $\underset{\sim}{x}_j$  for each sample, and the model is  $\underset{\sim}{y}_j = a + b \underset{\sim}{x}_j + \underset{\sim}{\varepsilon}_j$  for  $j = 1, \dots, k$ . In contrast, there are only different vectors of response variable  $\underset{\sim}{y}_j$  in LP, and the same vector of  $\underset{\sim}{x}$  is used as it is fixed in (3.1).

In this paper, our objective is to monitor the profile coefficients  $a$  and  $b$  of a process over time, rather than focusing on the process means. We aim to observe potential shifts resulting from changes in the coefficients.

---

### 4. Calculation of Profile Coefficients

---

This section introduces a novel nonparametric method for determining profile coefficients and the corresponding CC. This method does not require any fundamental assumptions about the data, which is an advantage of the entropy principle. We compare the results of our distribution-free method with the coefficients from  $LR$ . We refer to it as distribution-free because we do not make any assumptions about the distributions of  $X$  and  $Y$ . We define a two-dimensional vector  $(a, b)$ , then apply the  $T^2$ -Hotelling statistic to reduce the dimensionality to 1. Let  $\underset{\sim}{x} = (x_1, \dots, x_n)$  represents the fixed observations of  $X$  and let  $k$  denotes the sample numbers of  $Y$ , each with a length of  $n$ :

$$(\underset{\sim}{y}_1, \underset{\sim}{x}), \dots, (\underset{\sim}{y}_k, \underset{\sim}{x}).$$

Currently, we calculate  $(a, b)$  for  $k$  samples via  $ME$  and  $LR$ . Hence, we desire to present some notations first:

$$(4.1a) \quad \underset{\sim}{m}_1 = (\hat{a}_{1-ME}, \hat{b}_{1-ME}), \dots, \underset{\sim}{m}_k = (\hat{a}_{k-ME}, \hat{b}_{k-ME}),$$

$$(4.1b) \quad \underset{\sim}{l}_1 = (\hat{a}_{1-LR}, \hat{b}_{1-LR}), \dots, \underset{\sim}{l}_k = (\hat{a}_{k-LR}, \hat{b}_{k-LR}).$$

Vector  $(\widehat{a}_{1-ME}, \widehat{b}_{1-ME})$  includes the estimated values of coefficients in (3.1) using  $ME$  for the first sample named as  $m_1$ . The equivalent meaning is coming from  $(\widehat{a}_{1-LR}, \widehat{b}_{1-LR})$  that is called  $l_1$  with  $LR$  approximation. In the next step, we estimate the unknown distribution of each sample to calculate the corresponding coefficients  $m_1, \dots, m_k$ . As mentioned in section 2, some constraints are necessary to compute  $ME$  distribution. We choose six conditions presented as follows for  $j = 1, \dots, k$ :

$$(4.2) \quad \begin{cases} \int \int_{\mathbb{S}(X, Y_j)} f_{X, Y_j}(x, y) dx dy = 1, \\ \int \int_{\mathbb{S}(X, Y_j)} x f_{X, Y_j}(x, y) dx dy = \bar{x}, \\ \int \int_{\mathbb{S}(X, Y_j)} y f_{X, Y_j}(x, y) dx dy = \bar{y}_j, \\ \int \int_{\mathbb{S}(X, Y_j)} x^2 f_{X, Y_j}(x, y) dx dy = \bar{x}^2, \\ \int \int_{\mathbb{S}(X, Y_j)} y^2 f_{X, Y_j}(x, y) dx dy = \bar{y}_j^2, \\ \int \int_{\mathbb{S}(X, Y_j)} xy f_{X, Y_j}(x, y) dx dy = \bar{x}\bar{y}_j. \end{cases}$$

We determine the Lagrange coefficients  $\lambda_0, \dots, \lambda_5$  in such a way explained in (2.2) that the final  $f_{X, Y}(\cdot, \cdot)$  is a valid joint density function. To achieve this, we substitute the function from Equation (2.3) into the system of equations represented by (4.2). Subsequently, we solve the equations for the Lagrange coefficients. The estimated joint distribution functions of  $X$  and  $Y$  are obtained via  $ME$  for each sample  $(y_j, x)$ , where  $j = 1, \dots, k$ . As a result, we have  $k$  distributions corresponding to each sample. The parameters  $a$  and  $b$  in  $ME$  are defined as follows:

$$(4.3a) \quad \widehat{b}_{j-ME} = \frac{E_j[(X - E_j(X))(Y - E_j(Y_j))]}{E_j[(X - E_j(X))^2]},$$

$$(4.3b) \quad \widehat{a}_{j-ME} = E_j(Y_j) - \widehat{b}_{j-ME} E_j(X), \quad j = 1, \dots, k,$$

where  $E_j(\cdot)$  represents the mathematical expectation based on the  $ME$  density function, the index  $j$  corresponds to  $f_{X, Y_j}(x, y)$  for  $j = 1, \dots, k$ . The corresponding  $LR$  coefficients are computed for  $k$  samples:

$$(4.4a) \quad \widehat{b}_{j-LR} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_{ij} - \bar{Y}_j)}{\sum_{i=1}^n (X_i - \bar{X})^2},$$

$$(4.4b) \quad \widehat{a}_{j-LR} = \bar{Y}_j - \widehat{b}_{j-LR} \bar{X}, \quad j = 1, \dots, k.$$

For practical application, detailed information on calculating the coefficients is available in Subsection 5.1.

Our goal is to plot  $m_1, \dots, m_k$  and  $l_1, \dots, l_k$  to detect shifts of a process instead of monitoring its means. To do this, we use  $T^2$ -Hotelling statistic for all samples  $j = 1, \dots, k$  as below, and we name them  $T_{j-ME}^2$  and  $T_{j-LR}^2$  for  $ME$  and  $LR$ , respectively:

$$(4.5a) \quad T_{j-ME}^2 = (m_j - \bar{m})' S_m^{-1} (m_j - \bar{m}), \quad j = 1, \dots, k,$$

$$(4.5b) \quad T_{j-LR}^2 = (l_j - \bar{l})' S_l^{-1} (l_j - \bar{l}), \quad j = 1, \dots, k,$$

where  $\bar{m}$  and  $\bar{l}$  are two-dimension means related to  $m_1, \dots, m_k$  and  $l_1, \dots, l_k$ .  $S_m$  and  $S_l$  are the variance-covariance estimation matrices of  $m_1, \dots, m_k$  and  $l_1, \dots, l_k$ , respectively. We



define two Upper Control Limits ( $UCLs$ ) to assess the status of  $T^2$ . The first  $UCL$  used here is based on Fisher distribution, which is the same for  $T_{j-ME}^2$  and  $T_{j-LR}^2$ :

$$(4.6) \quad UCL_F = \frac{p(k+1)(k-1)}{k^2 - pk} F_{\alpha, p, k-p},$$

where  $p$  represents the number of profile coefficients, and in this case,  $p = 2$ . When the number of samples  $k$  exceeds 100, the control limits are adjusted to:

$$(4.7) \quad UCL_F = \frac{p(k-1)}{k-p} F_{\alpha, p, k-p}.$$

The second control limit is quantile of  $T_{1-ME}^2, \dots, T_{k-ME}^2$  and  $T_{1-LR}^2, \dots, T_{k-LR}^2$ . These control limits are calculated separately for  $ME$  and  $LR$ :

$$(4.8a) \quad UCL_{ME} = q_{ME},$$

$$(4.8b) \quad UCL_{LR} = q_{LR},$$

where  $q_{ME}$  and  $q_{LR}$  are  $(1 - \alpha)100\%$   $T^2$ -Hotelling quantile values of  $ME$  and  $LR$ , respectively.  $\alpha$  is the first type of error, sometimes determined as 0.05. We develop algorithms to comprehensively explain monitoring data using the  $ME$  and  $LR$  methods.

#### The $ME$ Algorithm:

1. Have  $k$  observation vectors of  $\tilde{y}$  and a vector  $\tilde{x}$  with  $n$  dimensions.
2. Calculate the unknown joint distributions of  $(\tilde{y}_j, \tilde{x})$  for  $j = 1, \dots, k$  via  $ME$ . To do that, substitute (2.3) in system (4.2) and solve it for  $\lambda$ 's.
3. Compute values of (4.3a) and (4.3b) for each  $\hat{b}_{j-ME}$  and  $\hat{a}_{j-ME}$  using the  $j$ th distribution. So, we have (4.1a).
4. Find (4.5a) values and plot the points.
5. Use  $UCLs$  (4.8a) along with either (4.6) or (4.7) to make a decision.

#### The $LR$ Algorithm:

1. Have  $k$  observation vectors of  $\tilde{y}$  and a vector  $\tilde{x}$  with  $n$  dimensions.
2. Calculate (4.4), so we have (4.1b).
3. Compute (4.5b) and plot the result points.
4. Use  $UCLs$  (4.8b) in conjunction with either (4.6) or (4.7) to derive the results for control situations.

These two methods are effective for detecting small shifts; however, the  $ME$  method is more effective. The values of  $ARL_0$  and  $ARL_1$  are estimated for stepwise shifts in intercepts and slopes in the simulation study section to support this assertion.

---

## 5. Simulation Studies

---

In this section, we employ presented methods to identify shifts in a comprehensive simulation study, as detailed in Subsection 5.2.  $ARL$  is the average number of IC samples plotted before detecting an OOC point. The  $ARL_0$  is typically calculated based on the first type of error, denoted as  $\alpha$ , for uncorrelated data, as shown below:

$$(5.1) \quad ARL_0 = \frac{1}{\alpha},$$

where  $\alpha$  represents the probability that an IC sample reaches the  $UCL$ . The term  $ARL_0$  is used to evaluate the performance of control limits, while  $ARL_1$  is applied to detect shifts in the process mean, referred to as OOC  $ARL$ , and defined as follows:

$$(5.2) \quad ARL_1 = \frac{1}{1 - \beta},$$

where  $\beta$  represents the second type of error, defined as the probability of remaining in the statistical control when a mean shift occurs. Typically, the control monitoring of a process comprises two phases: Phase *I* and Phase *II*. Appropriate control limits are established based on Phase *I*. Next, it is assessed whether the process is IC and if these established control limits are reliable for Phase *II* to detect unwanted changes. Furthermore, the parameters, including  $ARL_0$ , are calculated during this phase. The purpose is to promptly detect undesirable shifts or changes in the process parameters.

---

### 5.1. Method Explanations with a Simple Example

---

Let's examine a small simulation example and outline the steps for determining the nonparametric control limits using the methods described in this article. The simulated data is presented in Table 1. The applied constraints for the four samples in rows ( $j = 1, \dots, 4$ ) are:

$$(5.3) \quad \begin{cases} \int \int_{\mathbb{S}(X,Y_j)} f_{X,Y_j}(x,y) dx dy = 1, \\ \int \int_{\mathbb{S}(X,Y_j)} x f_{X,Y_j}(x,y) dx dy = \bar{x}, \\ \int \int_{\mathbb{S}(X,Y_j)} y f_{X,Y_j}(x,y) dx dy = \bar{y}_j, \\ \int \int_{\mathbb{S}(X,Y_j)} x y f_{X,Y_j}(x,y) dx dy = \overline{xy}_j. \end{cases}$$

Table 1: Simulated data in phase *I*.

sample		x				$\bar{y}$	$\overline{xy}$
		0.05	0.1	0.15	0.2		
$y$	1	0.135	1.434	1.228	2.133	1.233	0.190
	2	0.955	1.143	1.493	2.058	1.412	0.199
	3	0.267	1.122	1.350	1.948	1.172	0.179
	4	0.179	0.915	1.154	2.435	1.171	0.190

The final condition in (5.3) involves a straightforward dependency factor between the variables. The Lagrange function is written for each sample  $j = 1, \dots, 4$ :

$$\begin{aligned} L_j(f, \lambda_0, \lambda_1, \lambda_2, \lambda_3) = & - \int_0^{7.435} \int_0^5 \log(f_{X,Y_j}(x, y)) f_{X,Y_j}(x, y) dx dy \\ & - \lambda_0 \left\{ \int_0^{7.435} \int_0^5 f_{X,Y_j}(x, y) dx dy - 1 \right\} \\ & - \lambda_1 \left\{ \int_0^{7.435} \int_0^5 x f_{X,Y_j}(x, y) dx dy - \bar{x} \right\} \\ & - \lambda_2 \left\{ \int_0^{7.435} \int_0^5 y f_{X,Y_j}(x, y) dx dy - \bar{y}_j \right\} \\ & - \lambda_3 \left\{ \int_0^{7.435} \int_0^5 x y f_{X,Y_j}(x, y) dx dy - \overline{xy}_j \right\}. \end{aligned}$$

After taking the derivative with respect to  $f_{X,Y_j}(\cdot)$  and setting it equal to zero, like the steps in Section 2, we get the following function, which is similar to (2.3):

(5.4)

$$f_{X,Y_j}(x, y) = \exp(-1 - \lambda_{0j} - x\lambda_{1j} - y\lambda_{2j} - xy\lambda_{3j}), \quad j = 1, \dots, 4, \quad x \in (0, 5), \quad y \in (0, 7.435).$$

For simplicity, we nominate  $1 + \lambda_{0j}$  as  $\lambda_{0j}$  as a fixed term. The conditions are rewritten using function (5.4) for each  $j = 1, \dots, 4$ :

$$(5.5) \quad \begin{cases} \int_0^{7.435} \int_0^5 \exp(-\lambda_{0j} - x\lambda_{1j} - y\lambda_{2j} - xy\lambda_{3j}) dx dy = 1, \\ \int_0^{7.435} \int_0^5 x \exp(-\lambda_{0j} - x\lambda_{1j} - y\lambda_{2j} - xy\lambda_{3j}) dx dy = 0.125, \\ \int_0^{7.435} \int_0^5 y \exp(-\lambda_{0j} - x\lambda_{1j} - y\lambda_{2j} - xy\lambda_{3j}) dx dy = \bar{y}_j, \\ \int_0^{7.435} \int_0^5 xy \exp(-\lambda_{0j} - x\lambda_{1j} - y\lambda_{2j} - xy\lambda_{3j}) dx dy = \overline{xy}_j. \end{cases}$$

The system of equations (5.5) contains unknown parameters,  $\lambda_{0j}, \dots, \lambda_{3j}$ . *ME* coefficients are calculated by solving this system of equations for  $\lambda$ s. The number of samples is four, so system (5.5) must be solved four times to obtain the *ME* distributions. The Lagrange coefficients are presented in Table 2. For example, the approximated distribution of the first sample is:

$$(5.6) \quad f_{X,Y_1}(x, y) = \exp(2.064 - 9.558 x - 0.948 y + 1.024 x y), \quad x \in (0, 5), \quad y \in (0, 7.435).$$

Next, the profile coefficients must be computed using *ME* with the obtained distributions.

Table 2: Coefficients of *ME* for data in Table 1.

Sample	$\lambda_0$	$\lambda_1$	$\lambda_2$	$\lambda_3$
1	-2.064	9.558	0.948	-1.024
2	-1.844	9.046	0.776	-0.656
3	-2.106	9.467	0.992	-1.022
4	-2.147	9.760	1.018	-1.157

Equation (4.3) is used to determine  $\tilde{m}_1, \dots, \tilde{m}_k$ . For example, the procedure to obtain  $\tilde{m}_1$  is

as follows using (5.6):

$$\begin{aligned}
E_1(X) &= \int_0^{7.435} \int_0^5 x f_{X,Y_1}(x, y) dx dy = 0.125, \\
E_1[(X - E_1(X))^2] &= \int_0^{7.435} \int_0^5 (x - E_1(X))^2 f_{X,Y_1}(x, y) dx dy = 0.018, \\
E_1(Y_1) &= \int_0^{7.435} \int_0^5 y f_{X,Y_1}(x, y) dx dy = 1.233, \\
\hat{b}_{1-ME} &= \frac{1}{E_1[(X - E_1(X))^2]} \int_0^{7.435} \int_0^5 (x - E_1(X))(y - E_1(Y_1)) f_{X,Y_1}(x, y) dx dy \\
&= 2.041, \\
\hat{a}_{1-ME} &= E_1(Y_1) - \hat{b}_{1-ME} E_1(X) = 0.977,
\end{aligned}$$

The profile coefficients for all samples are computed using *ME* and *LR*, as shown in Table 3. The computation for the *LR* section of the table follows the traditional *LR* methodology. The values of  $T_{ME}^2$  are derived from Equation (4.5a).  $UCL_{ME}$ , is set at the 95th percentile of the  $T_{ME}^2$  values, which is equal to 1.938.  $UCL_{LR}$  and  $UCL_F$  are 2.006 and 71.25, respectively. The Fisher-based limit is illogical as it is excessively high. The same steps are applied to the data in phase *II* to determine  $T_{ME}^2$ . The difference between the calculations in phases *I* and *II* lies in the computation of the  $T^2$ -Hotelling statistic. In this case, the formulations (4.5a) and (4.5b) are utilized, incorporating  $\bar{m}$ ,  $S_m$ ,  $\bar{l}$ , and  $S_l$  from phase *I*, Montgomery (2019). This example aims to demonstrate the practical implementation of the algorithms. In the next subsection, a comprehensive example involving the *ARL* computation is presented.

Table 3: Profile coefficients via *ME* and *LR* with  $T^2$ -Hotelling.

Sample	<i>ME</i>			<i>LR</i>		
	$\hat{a}_{ME}$	$\hat{b}_{ME}$	$T_{ME}^2$	$\hat{a}_{LR}$	$\hat{b}_{LR}$	$T_{LR}^2$
1	0.977	2.041	0.405	-0.215	11.576	0.107
2	1.236	1.409	1.976	0.498	7.318	2.036
3	0.937	1.878	1.725	-0.146	10.542	1.471
4	0.890	2.249	1.248	-0.581	14.014	1.836

---

## 5.2. Comprehensive Simulation Study

---

To initiate the simulation study, we generate 100 samples with  $n = 5$  from the model  $Y = 2 + 3X + \varepsilon$ , where  $\varepsilon$  follows a normal distribution with a mean of 0 and a variance of 0.1. When the process is IC, the error must follow a zero-mean normal distribution. If it does not, the source of the error must be identified and eliminated. The fixed observations of  $X$  for all 100 samples are (2, 2.2, 2.4, 2.1, 2.7). We propose six constraints for this data, and the approximated density function using *ME* is:

$$(5.7) \quad f_{X,Y}(x, y) = \exp(-\lambda_0 - x\lambda_1 - y\lambda_2 - x^2\lambda_3 - y^2\lambda_4 - xy\lambda_5), \quad (x, y) \in \mathbb{S}(X, Y).$$

We calculate  $UCL$ s of the Fisher, 95% quantile in  $ME$ , and  $LR$ :

$$(5.8) \quad \begin{cases} UCL_F = 6.304, \\ UCL_{ME} = 5.591, \\ UCL_{LR} = 5.800. \end{cases}$$

To calculate  $ARL_0$  based on equation (5.1), we simulate 1000 batches of the correct model once more.  $ARL_0$  for  $ME$  are:

$$ARL_0 = \begin{cases} \frac{1}{0.059} = 16.949 \simeq 17, & UCL = UCL_F, \\ \frac{1}{0.08} = 12.500 \simeq 13, & UCL = UCL_{ME}, \end{cases}$$

and the related  $ARL_0$  for  $LR$  are:

$$ARL_0 = \begin{cases} \frac{1}{0.047} = 21.277 \simeq 22, & UCL = UCL_F, \\ \frac{1}{0.058} = 17.241 \simeq 18, & UCL = UCL_{LR}. \end{cases}$$

Keep in mind that the standard  $ARL_0$  for  $\alpha = 0.05$  is 20. The values of  $ARL_0$  indicate the number of samples required to trigger a false alarm when the process is IC. A higher number of  $ARL_0$  values signifies a better selection of control limits, resulting in false alarms occurring less frequently and later. Conversely, a lower  $ARL_0$  indicates accruing more false alarms. Therefore, the optimal choice of  $UCL$  is  $UCL_F$  in the  $LR$  context. We increase the quantile percentage to 98.75%, and the corresponding  $UCL$  values are:

$$(5.9) \quad \begin{cases} UCL_F = 9.354, \\ UCL_{ME} = 11.348, \\ UCL_{LR} = 6.905. \end{cases}$$

Then, we have:

$$ARL_0 = \begin{cases} \frac{1}{0.018} = 55.556 \simeq 56, & UCL = UCL_F, \\ \frac{1}{0.009} = 111.111 \simeq 112, & UCL = UCL_{ME}, \end{cases}$$

$$ARL_0 = \begin{cases} \frac{1}{0.01} = 100, & UCL = UCL_F, \\ \frac{1}{0.034} = 29.412 \simeq 30, & UCL = UCL_{LR}. \end{cases}$$

Interestingly, the results indicate a preference for  $ME$ , which exhibits the highest  $ARL_0$ . The standard  $ARL_0$  is 80 for  $\alpha = 0.0125$ . We simulate datasets using shifted models and testing methods to detect changes in the process means by  $ARL_1$ , as defined in (5.2). Three shifted models are presented that include intercepts, slopes, and a combination of both.

$$\begin{cases} \text{Model 1 : } Y = (2 + s) + 3X + \varepsilon, \\ \text{Model 2 : } Y = 2 + (3 + s)X + \varepsilon, \\ \text{Model 3 : } Y = (2 + s) + (3 + s)X + \varepsilon, \end{cases}$$

where shift  $s$  is from 0.01 to 0.32:

$$s = \{0.01, 0.02, 0.03, \dots, 0.3, 0.31, 0.32\}.$$



suggesting that  $LR$  computations are preferable. Thus, the primary aim of the  $ME$  method is to maintain high accuracy in the process; otherwise, the  $LR$  method is more suitable.

We provide the  $ARL_1$  for the models using equations (5.8) and (5.9) in Table 4, which is based on stepwise shifts. Table 4 consists of three sections: the intercept, slope, and mixed model. An additional column is included at the end to indicate which model has the minimum value of  $ARL_1$ . For small shifts, the  $ME$  using the  $UCL_{ME}$  exhibits the lowest value. However, as the shifts become more significant, the optimal control limit is the  $LR$  using the  $UCL_{LR}$ .

When the intercept shifts from the origin, the data experiences slight changes. However, when the slope changes, this alteration is amplified in the data, resulting in a more significant impact. These unwanted changes become more pronounced when both the intercept and slope are altered. Consequently, when only the intercept shifts, it becomes more challenging to detect. Referring back to Table 4, the  $ME$  can identify small changes more quickly than the  $LR$  method. If the technician's objective is not to detect minor changes, the complex computations involved in the entropy method can be bypassed in favor of using  $LR$ . The effectiveness of the entropy method is most evident when the goal is to identify small changes.

The  $ME$  can be applied in various manufacturing processes that should not experience minor shifts. These processes are particularly sensitive in critical productions, such as pharmaceutical production or high-cost procedures. Consequently, any potential changes are detected promptly to prevent financial losses.

---

## 6. Semiconductor Production Process

---

In this section, we work on a profile dataset from Zou et al. (2007), whose process is on deep reactive ion etching process (DRIE) from semiconductor manufacturing. They mentioned the source of used data in the DRIE process as one of the most important quality characteristics in the profile of a trench that may significantly impact the downstream operations, May et al. (1991). The desired profile is a smooth and vertical sidewall, as indicated in the centre sample in Figure 2. The first dataset included 18 samples with size  $n = 11$ , and the observed vector of independent variable is  $(-2.5, -2, -1.5, -1, -0.5, 0, 0.5, 1, 1.5, 2, 2.5)$ . The second set, for phase  $II$ , consisted of 14 samples of the same size.

We aim to check the second phase of the process to see if there are possible defective products. Zou et al. (2007) noted that the last sample of Phase  $II$  is OOC. Thus, our control limits should detect these changes in the process. We calculate the  $UCL_F$  and compare it with  $UCL_{ME}$  and  $UCL_{LR}$ . We compute  $ME$  distributions of all 18 samples considering (5.7) to find the  $UCL_{ME}$ . We provide six Lagrange coefficients of  $ME$  according to constraints (4.2) in Table 5. Profile coefficients of all 18  $ME$  distributions are in Table 6 with  $T^2$ -Hotelling values and  $LR$  results. The next step is to specify beneficial  $UCLs$  to monitor the rest of the process:

$$(6.1) \quad \begin{cases} UCL_F = 8.151, \\ UCL_{ME} = 4.857, \\ UCL_{LR} = 4.857. \end{cases}$$

Table 4:  $ARL_1$ 's for 95% and 98.75%.

Quantile	95%				98.75%				Minimum $ARL_1 \downarrow$
Methods	$LR$		ME		$LR$		ME		
$UCL$	$UCL_F$	$UCL_{LR}$	$UCL_F$	$UCL_{ME}$	$UCL_F$	$UCL_{LR}$	$UCL_F$	$UCL_{ME}$	
$ARL_0$	22	18	17	13	100	30	56	112	$\leftarrow$ Maximum $ARL_0$
									$UCL_{ME} - 98.75\%$
Shifts	Intercept								
0.01	22.727	17.241	17.543	12.500	83.333	27.027	43.478	62.5	$UCL_{ME} - 95\%$
0.02	22.727	18.182	14.925	11.363	76.923	29.412	35.714	66.667	$UCL_{ME} - 95\%$
0.03	13.889	11.765	10.101	8.620	38.465	17.241	18.182	25.641	$UCL_{ME} - 95\%$
0.04	10.870	9.009	10.989	8.333	35.714	14.286	21.739	40.000	$UCL_{ME} - 95\%$
0.05	6.897	5.988	9.433	7.299	21.739	8.333	20.408	32.258	$UCL_{LR} - 95\%$
0.06	5.814	4.902	7.299	6.060	16.393	7.092	18.519	31.25	$UCL_{LR} - 95\%$
0.07	4.854	4.274	6.896	5.847	11.494	5.988	14.085	23.256	$UCL_{LR} - 95\%$
0.08	3.268	3.003	6.060	4.784	6.803	3.717	12.346	19.231	$UCL_{LR} - 95\%$
0.09	2.747	2.469	5.000	3.861	5.263	3.067	10.526	17.544	$UCL_{LR} - 95\%$
0.10	2.132	1.953	4.424	3.484	3.817	2.381	10.101	15.625	$UCL_{LR} - 95\%$
0.15	1.217	1.185	1.745	1.557	1.538	1.242	3.322	4.950	$UCL_{LR} - 95\%$
0.20	1.012	1.009	1.165	1.119	1.063	1.018	1.575	1.965	$UCL_{LR} - 95\%$
0.30	1.000	1.000	1.000	1.000	1.000	1.000	1.007	1.022	-
Shifts	Slope								
0.01	26.316	19.231	13.888	10.526	111.111	32.258	43.478	76.923	$UCL_{ME} - 95\%$
0.02	9.709	8.403	9.174	7.299	29.412	12.658	18.182	29.412	$UCL_{ME} - 95\%$
0.04	2.703	4.464	4.901	3.676	5.236	3.049	10.753	18.182	$LR - UCL_F - 95\%$
0.06	1.374	1.328	2.192	1.886	1.980	1.449	4.202	6.803	$UCL_{LR} - 95\%$
0.07	1.136	1.115	1.485	1.353	1.395	1.188	2.358	3.289	$UCL_{LR} - 95\%$
0.09	1.014	1.010	1.081	1.053	1.057	1.022	1.325	1.639	$UCL_{LR} - 95\%$
0.10	1.001	1.001	1.033	1.015	1.008	1.003	1.178	1.348	$UCL_{LR} - 95\%$
Shifts	Intercept and Slope								
0.01	15.873	12.346	12.048	10.101	52.632	18.182	27.778	40.000	$UCL_{ME} - 95\%$
0.02	5.051	4.329	7.751	6.211	11.905	5.917	18.182	31.250	$UCL_{LR} - 95\%$



0.03	2.375	2.119	3.937	3.174	4.386	2.681	7.813	12.346	$UCL_{LR} - 95\%$
0.04	1.412	1.328	2.192	1.912	2.004	1.490	4.255	6.410	$UCL_{LR} - 95\%$
0.05	1.139	1.120	1.445	1.344	1.342	1.166	2.342	3.546	$UCL_{LR} - 95\%$
0.06	1.026	1.021	1.160	1.111	1.073	1.037	1.529	2.045	$UCL_{LR} - 95\%$
0.07	1.007	1.005	1.031	1.018	1.017	1.008	1.183	1.362	$UCL_{LR} - 95\%$
0.08	1.000	1.000	1.008	1.003	1.002	1.000	1.037	1.093	-
0.09	1.000	1.000	1.001	1.000	1.000	1.000	1.012	1.030	-
0.10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.004	-

We compute  $ME$  Lagrange coefficients of the 14 samples of phase  $II$  in Table 7 using (5.7). The corresponding  $ME$  and  $LR$  profile coefficients, along with the  $T^2$ -Hotelling statistics, are presented in Table 8. All coefficients and statistics are identical for the  $ME$  and  $LR$ . Both methods identify the 14th sample as OOC, as highlighted in Table 8. The intriguing part of the result is that the traditional Fisher CC cannot detect this shift. Back to Equation (6.1), the proposed charts have the same  $UCL$ , but the Fisher-based has a higher  $UCL$ , which makes the superiority of  $ME$  and  $LR$  performance. As a consequence, since  $ME$  and  $LR$  results are the same, the  $LR$  is preferable due to its simpler computations. In the next section, we tackle another example that shows the overall superiority of  $ME$ .

Table 5: The Lagrange coefficients of  $ME$  for IC DRIE process.

Number	$\lambda_0$	$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$	$\lambda_5$
1	2.148	0.087	-0.734	0.204	0.258	-0.061
2	2.154	-0.014	-0.752	0.200	0.263	0.010
3	2.150	-0.072	-0.726	0.203	0.255	0.051
4	2.275	-0.043	-0.818	0.201	0.258	0.027
5	2.127	0.021	-0.847	0.200	0.307	-0.015
6	2.293	0.024	-0.808	0.200	0.251	-0.015
7	2.082	0.064	-0.563	0.203	0.216	-0.049
8	2.340	0.098	-0.907	0.203	0.276	-0.059
9	2.265	-0.090	-0.703	0.204	0.221	0.056
10	2.202	-0.013	-0.754	0.200	0.252	0.008
11	2.196	-0.163	-0.852	0.211	0.290	0.111
12	2.299	0.086	-0.799	0.203	0.246	-0.053
13	2.397	0.053	-0.812	0.201	0.231	-0.030
14	2.307	-0.005	-0.997	0.200	0.317	0.003
15	2.248	-0.027	-0.563	0.200	0.180	0.017
16	2.379	-0.044	-0.739	0.201	0.210	0.025
17	2.565	-0.075	-1.139	0.201	0.313	0.041
18	2.104	0.001	-0.605	0.200	0.225	-0.001

Table 6:  $ME$  and  $LR$  coefficients with  $T^2$ -Hotellings for IC data of DRIE process.

Number	$ME$			$LR$		
	$\hat{a}_{ME}$	$\hat{b}_{ME}$	$T_{ME}^2$	$\hat{a}_{LR}$	$\hat{b}_{LR}$	$T_{LR}^2$
1	1.422	0.118	2.864	1.422	0.118	2.864
2	1.428	-0.018	0.700	1.428	-0.018	0.700
3	1.425	-0.099	1.985	1.425	-0.099	1.985
4	1.585	-0.053	0.376	1.585	-0.053	0.376
5	1.383	0.025	1.416	1.383	0.025	1.416
6	1.613	0.029	0.423	1.613	0.029	0.423
7	1.303	0.114	4.802	1.303	0.114	4.802
8	1.645	0.108	2.399	1.645	0.108	2.399
9	1.594	-0.128	2.166	1.594	-0.128	2.166
10	1.495	-0.017	0.138	1.495	-0.017	0.138
11	1.471	-0.192	5.170	1.471	-0.192	5.170
12	1.623	0.107	2.161	1.623	0.107	2.161
13	1.760	0.065	3.098	1.760	0.065	3.098

14	1.572	-0.005	0.037	1.572	-0.005	0.037
15	1.565	-0.047	0.250	1.565	-0.047	0.250
16	1.758	-0.060	2.609	1.758	-0.060	2.606
17	1.818	-0.065	4.121	1.818	-0.065	4.121
18	1.347	0.002	1.936	1.347	0.002	1.936

Table 7: The Lagrange coefficients of  $ME$  for phase  $II$  of DRIE process.

Number	$\lambda_0$	$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$	$\lambda_5$
1	2.351	0.094	-1.051	0.203	0.327	-0.059
2	2.266	-0.109	-0.814	0.205	0.259	0.069
3	2.237	-0.018	-0.720	0.200	0.232	0.012
4	2.218	0.1001	-0.799	0.204	0.265	-0.066
5	2.228	0.0389	-0.702	0.201	0.228	-0.025
6	2.173	0.0480	-0.635	0.201	0.218	-0.033
7	2.247	-0.000	-0.562	0.200	0.180	0.000
8	2.211	-0.020	-0.638	0.200	0.211	0.013
9	2.274	-0.077	-0.589	0.203	0.183	0.048
10	2.277	0.011	-0.738	0.200	0.230	-0.007
11	2.140	0.051	-0.611	0.202	0.218	-0.037
12	2.268	0.019	-0.563	0.200	0.176	-0.012
13	2.322	0.019	-0.787	0.200	0.237	-0.012
14	2.154	-0.08	-0.470	0.205	0.171	0.061

---

## 7. Pharmacy Profile Data

---

In the previous example, both charts act the same, but this example shows the preference of  $ME$ . We choose a pharmaceutical industry profile dataset adapted from [Shah et al. \(1998\)](#). In the pharmaceutical industry, many processes are under restricted statistical control. For instance, in the production of a particular tablet, many characteristics are observed to ensure that the product is of a standard quality. [Ma et al. \(2000\)](#) indicate that the U.S. food and drug administration had to define many conductance of scale-up and post-approval changes for different types of shifts during production processes. In this regard, the statistical quality CCs are influential in pharmaceutical processes.

There are six groups of profile data in [Shah et al. \(1998\)](#), each containing the response variable in four columns as the cumulative dissolution of tablets at 30, 60, 90, and 180 minutes. The sample sizes have to be 12 according to the guidance of "Dissolution Testing of Immediate Release Solid Oral Dosage Forms", [Ma et al. \(2000\)](#). The first group is a pre-change as a dissolve reference in different minutes. The other five batches are from post-change processes. The beneficial aspect is detecting the similarity or dissimilarity of 5 batches from the reference group. For  $ME$ , we applied model (5.7) for the unknown approximation of the density function. Then, we calculate  $T^2$ -Hotelling statistics for these profiles via  $ME$

and  $LR$  in Table 9.  $UCLs$  are at the confidence level of 0.9973:

$$(7.1) \quad \begin{cases} UCL_F = 26.977, \\ UCL_{ME} = 6.813, \\ UCL_{LR} = 5.884. \end{cases}$$

Shah et al. (1998) declared that all batches have unwanted changes in the amount of dissolution. The dissolved values of batch 4 related to 30-minute are only dissimilar to the reference data in this group.  $ME$  reflects this reality via different  $T^2$ -Hotelling values in Table 9. Without any ambiguity, all batches except batch 4 are too different from the reference. Although all tablets of batch 4 are OOC based on  $UCL_{ME}$ , they are not as gigantic as other batch values. The larger the  $T^2$ -Hotelling values, the greater the difference compared to the reference batch. So,  $ME$  detects undesirable changes concerning the reference group. Nine tablets in batch 4 are IC according to the  $T_{ME}^2$  and  $UCL_F$ , which means that the traditional way cannot realize changes. The result of  $LR$  in Table 9 is quite different.  $LR$  does not completely detect changes in batch 1, and there are IC tablets in bold. Also, 29 IC samples exist related to  $UCL_F$  pointed by stars. Moreover, some tablets in batch 1 are IC, which is far from the reality.  $LR$  decision has some shortcomings. Thus, the  $ME$  power is straightforward, and other methods cannot indicate the post-change of the dataset in batches.

We have other information from Shah et al. (1998). The dissolved differences of batch 2 are 15% more than the reference group at 30-minute, but the differences are reduced to less than 8% at 60, 90, and 180-minutes. So, one of the diversity of batch 2 comes from 30-minute data. We recalculate in the absence of 30-minute data in Table 10. Batch 4 at 30-minute differs from the reference group more than 12%. The 30-minute data are excluded, and the corresponding Hotelling statistic is recalculated in Table 10. The confidence level of these  $UCLs$  are 0.9973%:

$$(7.2) \quad \begin{cases} UCL_F = 26.977, \\ UCL_{ME} = 7.895, \\ UCL_{LR} = 7.233. \end{cases}$$

According to Table 10, in  $UCL$  (7.2) columns,  $T^2$ -Hotelling statistics are reduced for batches 2 and 4, which means that their discrepancies are diminished from the reference group. Batch 2 is still OOC, and  $ME$  reflects it obviously, but batch 4 becomes more similar to the reference mentioned in bold. The conclusion of  $LR$  shown in Table 10 is the same. All tablets are OOC for batch 2. Almost all of them for batch 4 are IC, the same as the  $UCL_{ME}$ . Shah et al. (1998) expressed about batch 3 that the differences are more than 12% at 90-minute between this batch and the reference group, and the differences become less than 10% in the absence of data for 90-minute. So, we remove the 90-minute data and recalculate it all. The related  $T^2$ -Hotelling is in Table 10 with their related reference amounts. The used  $UCLs$  are at the confidence level 0.9973:

$$(7.3) \quad \begin{cases} UCL_F = 26.977, \\ UCL_{ME} = 6.407, \\ UCL_{LR} = 5.865. \end{cases}$$

Table 8: ME and LR coefficients for DRIE process in phase II with  $T^2$ -Hotelling.

Number	ME			LR		
	$\hat{a}_{ME}$	$\hat{b}_{ME}$	$T_{ME}^2$	$\hat{a}_{LR}$	$\hat{b}_{LR}$	$T_{LR}^2$
1	1.605	0.090	1.511	1.605	0.090	1.511
2	1.574	-0.134	2.303	1.574	-0.134	2.303
3	1.550	-0.025	0.052	1.550	-0.025	0.052
4	1.510	0.125	2.482	1.510	0.125	2.482
5	1.537	0.055	0.540	1.537	0.055	0.540
6	1.456	0.075	1.297	1.456	0.075	1.297
7	1.564	-0.001	0.023	1.564	-0.001	0.023
8	1.515	-0.031	0.137	1.515	-0.031	0.137
9	1.611	-0.131	2.347	1.611	-0.131	2.347
10	1.605	0.015	0.256	1.605	0.015	0.256
11	1.403	0.084	2.091	1.403	0.084	2.091
12	1.598	0.033	0.374	1.598	0.033	0.374
13	1.659	0.025	0.801	1.659	0.025	0.801
14	1.371	-0.178	<b>5.773</b>	1.371	-0.178	<b>5.773</b>

Table 9:  $T^2$ -Hotelling of tablets in phases I and II. The asterisk (\*) indicates the IC values concerning  $UCL_F$ , while the bold values pertain to the IC samples using  $UCL_{ME}$  and  $UCL_{LR}$ . The utilized control limits are in (7.1).

Number	Reference	Batch 1	Batch 2	Batch 3	Batch 4	Batch 5
ME						
1	0.910	2531250	40212322	24964618	16.395*	18160920
2	0.426	2528833	40139281	24937534	19.672*	18175083
3	0.942	2524719	40201702	24942289	20.274*	18143501
4	1.633	2524318	40165632	24924941	13.657*	18140501
5	6.867	2507804	40231679	24950972	38.401	18161265
6	0.705	2503931	40195131	24925518	20.163*	18155740
7	1.090	2525108	40156018	24964433	22.483*	18159950
8	2.198	2524320	40204586	24949374	17.023*	18157899
9	0.374	2527543	40223125	24961179	26.166*	18151208
10	1.313	2529049	40190721	24955811	23.062*	18171547
11	5.041	2523157	40213998	24946082	28.308	18163695
12	2.083	2512610	40188450	24942731	37.346	18160788
LR						
1	0.754	<b>3.606*</b>	9.745*	22.387*	26.566*	30.804
2	0.546	<b>1.257*</b>	66.229	16.150*	35.799	30.211
3	1.066	<b>4.799*</b>	10.038*	16.779*	46.385	55.707
4	0.265	<b>3.286*</b>	50.992	37.221	23.326*	53.688
5	5.944	63.696	7.94085*	20.633*	43.577	25.901*
6	1.112	77.022	23.14788*	11.417*	48.051	41.270
7	1.131	10.347*	49.88452	25.844*	42.771	36.815
8	1.893	<b>5.028*</b>	18.980*	19.095*	29.292	42.0778
9	0.490	<b>0.781*</b>	6.022*	27.745	39.823	46.064
10	1.574	<b>2.557*</b>	28.837	23.897*	39.109	24.267*
11	3.910	11.328*	15.905*	7.477*	40.167	27.370
12	0.791	38.186	32.921	45.483	59.874	36.490

Table 10:  $T^2$ -Hotelling in the absence of some data. The asterisk (\*) values mean IC samples concerning  $UCL_F$ , and the bold ones are the IC samples according to  $UCL_{ME}$  and  $UCL_{LR}$ .

Number	$UCL$ (7.2)			$UCL$ (7.3)		$UCL$ (7.4)	
	Reference except 30- min	Batch 2 except 30-min	Batch 4 except 30-min	Reference except 90-min	Batch 3 except 90-min	Reference except 60-min	Batch 5 except 60-min
ME							
1	0.860	32054088	<b>0.860*</b>	0.646	11342273	1.031	7100891
2	0.097	31966736	<b>0.097*</b>	0.662	11326617	0.333	7108760
3	1.074	32042515	<b>1.074*</b>	0.765	11329719	0.820	7094211
4	7.932	32003090	7.932*	0.989	11320807	0.752	7094148
5	6.674	32066905	<b>6.674*</b>	6.483	11335365	6.522	7095663
6	1.510	32024502	<b>1.510*</b>	0.674	11318497	0.699	7099653
7	0.479	31984663	<b>0.480*</b>	0.974	11343690	1.517	7102175
8	2.329	32042369	<b>2.329*</b>	3.084	11334122	2.885	7102262
9	0.370	32054088	<b>0.370*</b>	0.369	11341106	0.414	7097997
10	0.250	32041637	<b>0.250*</b>	1.375	11338460	1.626	7104667
11	5.784	32055156	<b>5.784*</b>	3.937	11330098	2.434	7100702
12	4.926	32031147	<b>4.926*</b>	1.024	11331677	0.892	7102083
LR							
1	0.781	19.479*	<b>0.781*</b>	0.745	18.700*	0.811	9.815*
2	0.104	31.607	<b>0.104*</b>	0.618	12.680*	0.414	14.402*
3	1.064	10.549*	<b>1.064*</b>	0.814	15.636*	1.060	23.690*
4	7.280	14.664*	7.280*	0.341	29.589	0.239	19.501*
5	5.696	31.996	<b>5.696*</b>	5.938	19.4079*	5.935	8.271*
6	1.456	10.165*	<b>1.456*</b>	0.947	8.489*	1.052	15.840*
7	0.347	17.764*	<b>0.349*</b>	1.122	25.023*	1.404	14.091*
8	2.091	16.018*	<b>2.091*</b>	2.987	17.193*	2.456	17.012*
9	0.505	19.480*	<b>0.505*</b>	0.435	23.507*	0.548	18.130*
10	0.336	9.787*	<b>0.336*</b>	1.513	22.448*	1.632	8.317*
11	5.169	21.787*	<b>5.170*</b>	3.465	6.065*	2.043	6.871*
12	3.795	10.816*	<b>3.795*</b>	0.593	38.090	0.337	18.874*

The proportion of the reference and batch 3 in the absence of 90-minute data is strongly rejected by  $ME$  using  $UCL_F$  and  $UCL_{ME}$  (7.3). Although  $T^2$ -Hotelling values are too large here, they are less than the calculated values in Table 9 related to batch 3. The result for  $LR$  is the same, and all dissolutions are OOC, which declares that batch 3 is not equivalent to the reference. We come across a very different result with  $T^2_{LR}$  and  $UCL_F$ . Almost all of these samples become IC, which is a massive disparity in decision-making. Further information is that the difference between the batch 5 and the reference is more than 17% for dissolved amounts of tablets at 60-minute data. It gets less than 10% when the 60-minute data are absent. The  $T^2$ -Hotelling values are recalculated in Table 10.  $UCLs$  at the significant level of 0.0027 are:

$$(7.4) \quad \begin{cases} UCL_F = 26.977, \\ UCL_{ME} = 6.414, \\ UCL_{LR} = 5.831. \end{cases}$$

The likeness between batch 5 and the reference is denied by both methods in the absence of 60-minute data. Although all  $T^2$ -Hotelling amounts via  $ME$  are less than the corresponding ones in Table 9, they show the existing differences. This decreasing in values means the reduction of perturbation in the samples that happened by removing the 60-minute. This pharmaceutical example demonstrates the superiority of the  $ME$  method, which makes decisive decisions,

while the *LR* method is unable to distinguish changes properly. Therefore, when the process produces a vital product, quality control must be reliable enough.

---

## 8. Conclusion

---

In statistical quality CCs, the purpose is to find suitable limits to check processes during time to keep them IC or to detect shifts as soon as possible to save money and time. There are some methods for this aim, and most charts are according to process means and normal distribution. To our knowledge, they are insufficient to detect different kinds of shifts. Therefore, we present a profile investigation of change points according to their coefficients. LP models are applied to implement our new nonparametric chart. We consider two estimation methods. The first method is an entropy-based distribution-free novel CC, compared with *LR*. We use a  $T^2$ -Hotelling statistic to handle vector dimensions.

We design a CC to detect changes in production processes by monitoring coefficients rather than means. We define three shifted models and calculate  $\beta$ ,  $ARL_0$ , and  $ARL_1$  to assess the chart performance. We conclude from the simulation results that both methods can detect small shifts; however, the *ME* method demonstrates superior performance with fewer false errors. In conclusion, we present two real datasets that contain change points within the samples. The small shift is not discernible when analyzing mean CCs; however, changes can be easily detected by *ME* and *LR* using profile data. Eventually, we present the performance of charts in three scenarios: a simulation example, semiconductor production, and pharmaceutical profiles. The effectiveness of the *ME* method in the second example is relatively outstanding.

---

## Data Availability Statement

---

The data that support the findings of sections 6 and 7 are openly available in Zou et al. (2007) (Zou et al., 2007) at <https://doi.org/10.1198/004017007000000164> and Shah et al. (1998) (Shah et al., 1998) at <https://doi.org/10.1023/A:1011976615750>, respectively.

---

## Declarations

---

This work was supported by the Innovative Research Group Project of the National Natural Science Foundation of China (Grant No. 51975253). The authors report no conflicts of interest or personal ties that may affect the work in this paper. The contributions are as follows:

Seyedeh Azadeh Fallah Mortezaejad: Methodology, Conceptualization, Validation, Writing-original draft, Writing-review and editing.

Ruo Chen Wang: Supervision, Methodology, Investigation, Writing- original draft, Writing-review and editing.

Gholamreza Mohtashami Borzadaran: Formatting, review and editing.  
 Renkai Ding: Formatting, review and editing.  
 Kim Phuc Tran: Conceptualization, Investigation, Writing- review and editing.

---

## ACKNOWLEDGMENTS

---

To the memory of Prof Bahram Sadeghpour Gildeh.

---

## REFERENCES

---

- Che, Y., Deng, Y., and Yuan, Y.-H. (2022). Maximum-entropy-based decision-making trial and evaluation laboratory and its application in emergency management. *Journal of Organizational and End User Computing (JOEUC)*, 34(7):1–16.
- Dalleiger, S. and Vreeken, J. (2020). The relaxed maximum entropy distribution and its application to pattern discovery. In *2020 IEEE International Conference on Data Mining (ICDM)*, pages 978–983. IEEE.
- Geng, W., Haruna, S. A., Li, H., Kademi, H. I., and Chen, Q. (2023). A novel colorimetric sensor array coupled multivariate calibration analysis for predicting freshness in chicken meat: A comparison of linear and nonlinear regression algorithms. *Foods*, 12(4):720.
- Imran, M., Dai, H.-L., Zaidi, F. S., Tran, K. P., Abbas, Z., and Nazir, H. Z. (2023). Incorporating principal component analysis into hotelling t2 control chart for compositional data monitoring. *Computers & Industrial Engineering*, 186:109755.
- Jaynes, E. T. (1957). Information theory and statistical mechanics. *Physical review*, 106(4):620.
- Kopp, R. E. (1962). Pontryagin maximum principle. In *Mathematics in Science and Engineering*, volume 5, pages 255–279. Elsevier.
- Ma, M.-C., Wang, B. B., Liu, J.-P., and Tsong, Y. (2000). Assessment of similarity between dissolution profiles. *Journal of Biopharmaceutical Statistics*, 10(2):229–249.
- Macedo, P. (2024). A two-stage maximum entropy approach for time series regression. *Communications in Statistics-Simulation and Computation*, 53(1):518–528.
- May, G. S., Huang, J., and Spanos, C. J. (1991). Statistical experimental design in plasma etch modeling. *IEEE Transactions on Semiconductor Manufacturing*, 4(2):83–98.
- Montgomery, D. C. (2019). *Introduction to statistical quality control*. John wiley & sons.
- Nancy, M., Joshi, H., and Dhandra, B. (2023). Regression control charts-a survey. *Journal of Pharmaceutical Negative Results*, pages 1078–1086.
- Perdikis, T., Celano, G., and Chakraborti, S. (2023). Distribution-free control charts for monitoring scale in finite horizon productions. *European Journal of Operational Research*.
- Roshanbin, N., Ershadi, M., and Niaki, S. (2022). Multi-objective economic-statistical design of simple linear profiles using a combination of nsga-ii, rsm, and topsis. *Communications in Statistics-Simulation and Computation*, 51(4):1704–1720.



- Shah, V. P., Tsong, Y., Sathe, P., and Liu, J.-P. (1998). In vitro dissolution profile comparison—statistics and analysis of the similarity factor,  $f_2$ . *Pharmaceutical research*, 15:889–896.
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, 27(3):379–423.
- Wibawati, W., Ahsan, M., Khusna, H., Qori'atunnadyah, M., and Udiatami, W. M. (2022). Multivariate control chart based on neutrosophic hotelling  $t_2$  statistics and its application. *JTAM (Jurnal Teori dan Aplikasi Matematika)*, 6(1):85–92.
- Xie, M., Kong, H., and Goh, T. (2000). Exponential approximation for maintained weibull distributed component. *Journal of Quality in Maintenance Engineering*, 6(4):260–269.
- Yeganeh, A. and Shongwe, S. C. (2023). A novel application of statistical process control charts in financial market surveillance with the idea of profile monitoring. *PloS one*, 18(7):e0288627.
- Zaidi, F. S., Dai, H.-L., Imran, M., and Tran, K. P. (2023). Analyzing abnormal pattern of hotelling  $t_2$  control chart for compositional data using artificial neural networks. *Computers & Industrial Engineering*, 180:109254.
- Zhao, Z., Li, H., Liu, J., and Yang, S. X. (2020). Control method of seedbed compactness based on fragment soil compaction dynamic characteristics. *Soil and Tillage Research*, 198:104551.
- Zou, C., Tsung, F., and Wang, Z. (2007). Monitoring general linear profiles using multivariate exponentially weighted moving average schemes. *Technometrics*, 49(4):395–408.