



HAL
open science

On Comparison of Multiple Non-linear Profiles for Process Selection

Chen-Ju Lin, Pei-Ying Lin

► **To cite this version:**

Chen-Ju Lin, Pei-Ying Lin. On Comparison of Multiple Non-linear Profiles for Process Selection. IFIP International Conference on Advances in Production Management Systems (APMS), Aug 2018, Seoul, South Korea. pp.392-398, 10.1007/978-3-319-99704-9_48 . hal-02164904

HAL Id: hal-02164904

<https://inria.hal.science/hal-02164904>

Submitted on 25 Jun 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

On Comparison of Multiple Non-Linear Profiles for Process Selection

Chen-ju Lin^{1*} and Pei-Ying Lin¹

¹ Yuan Ze University, Taoyuan 32003, Taiwan
Chenju.lin@saturn.yzu.edu.tw

Abstract. A product can be produced by using alternative processes. The processes that can produce the product having consistent quality performance as the existing process are qualified alternative processes. Having several qualified alternative processes can relax the conditions for production scheduling and increase production capacity. This research focuses on analyzing the quality characteristic which is nonlinear profile. The quality characteristic, which is the response variable, has a non-linear functional relationship with the explanatory variable. The objective of this research is to select a subset of qualified alternative processes among I alternative processes, $I > 1$. A Step-up Test Procedure is proposed to compare the non-linear profile of each process with the existing process. Polynomial regression is used to estimate the non-linear profile models before applying the step-up sequential tests. The proposed method can effectively test the significance of the differences among processes, while controlling the overall error rate of testing I processes below α .

Keywords: Non-linear Profile, Multiple Comparisons, Sequential Tests.

1 Introduction

A product can be manufactured or assembled by using different processes. Having multiple alternative processes in production can increase utilization, reduce production time, or decrease production cost. However, the product quality under alternative processes may differ. Process Selection is a critical problem to production management. The objective of the problem is to select the qualified alternative processes that can produce the required specifications while having consistent quality performance as the existing process. This research aims at tackling the Process Selection problem under non-linear profile data, where there is a non-linear functional relationship between the response variable and the explanatory variable.

Most literature monitored profile data obtained from one manufacturing process. The monitoring process signals once the process is out-of-control [1-3]. However, the existing methods cannot be applied to solve the Process Selection problem where several processes are simultaneously compared with the existing process. The existing methods did not consider the overall error rate of conducting multiple comparisons. Conse-

quently, we propose the Step-up Test Procedure to select the qualified alternative processes from I processes under non-linear profile data, $I > 1$. The proposed method considers data variation from different alternative processes, and controls the overall error rate of conducting many-to-one comparisons below α .

The following paper is organized as follows. Section 2 reviews the related literature about profile monitoring. Section 3 introduces the proposed Step-up Test Procedure for the Process Selection problem. Section 4 presents the simulation results followed by the conclusions in Section 5.

2 Literature Review

Quality characteristics often appear to be non-linear profile data in industry [4, 5]. Polynomial regression models were commonly used to fit non-linear profile data [6, 7]. After model estimation, researchers used multiple control charts to monitor the coefficients and the variance of the estimated models. [8, 9] applied sine and hyperbolic tangent functions for data fitting before using the Hotelling's T^2 control chart for monitoring. Other researchers used wavelet to transform non-linear data [10, 11]. Their studies applied the T^2 or the Cumulative Sum (CUSUM) control charts for detecting mean shifts in non-linear profile data. The CUSUM control chart was shown to be more sensitive to small shifts than the T^2 chart.

To select a subset of qualified alternative processes, many-to-one comparisons need to be made. The aforementioned methods aim at monitoring one process, and signal when the process is out-of-control. The existing methods cannot be directly applied to solve the Process Selection problem which compares multiple processes. Consequently, we propose the Step-up Test Procedure to tackle the Process Selection problem under non-linear profile data. The method selects a subset of qualified alternative processes from I candidate processes, $I > 1$.

3 Methodology

The proposed method tests the discrepancy between the non-linear profiles of each alternative process and the existing process, indexed as 0. Let the quality characteristic (response variable) be a non-linear function of the explanatory variable, $y_{i,j} = f_i(x_{i,j}) + \varepsilon_{i,j}$. Let $x_{i,j}$, $y_{i,j}$, and $\varepsilon_{i,j}$ be the j th explanatory variable, response variable, and the error term of process i , respectively, $i = 0, \dots, I, j = 1, \dots, n$. Assume that $\varepsilon_{i,j}$ is normally distributed with a mean 0 and a standard deviation σ_i , $i = 0, \dots, I$. The explanatory variable is bounded for practical usage, $a < x_{i,j} < b$. The proposed method contains three steps: parameters estimation, test statistics calculation, and the step-up procedure, which are explained in the following subsections.

3.1 Parameters Estimation

Collect n samples from each process, $(x_{i,j}, y_{i,j})$, $i = 0, \dots, I, j = 1, \dots, n$. The non-linear profile data are fitted by using polynomial regression models, $y_{i,j} = \beta_{i,0} + \beta_{i,1} x_{i,j} + \dots +$

$\beta_{i,p}(x_{i,j})^p + \varepsilon_{i,j}$, $i = 0, \dots, I, j = 1, \dots, n$. The estimator of $\beta_i = [\beta_{i,0}, \beta_{i,1}, \dots, \beta_{i,p}]^T$ is calculated as $\hat{\beta}_i = (X_i^T X_i)^{-1} X_i^T y_i$ by using the least squares method, where

$$X_i = \begin{bmatrix} 1 & x_{i,1} & \cdots & x_{i,1}^p \\ \vdots & \vdots & \cdots & \vdots \\ 1 & x_{i,n} & \cdots & x_{i,n}^p \end{bmatrix} \quad (1)$$

and $y_i = [y_{i,1}, \dots, y_{i,n}]^T$, $i = 0, \dots, I$. The mean squared error is used to estimate variance

$$\hat{\sigma}_i^2 = \frac{(y_i - X_i \hat{\beta}_i)^T (y_i - X_i \hat{\beta}_i)}{n-p-1}, i = 0, \dots, I. \quad (2)$$

Since the highest power term of the polynomial regression model is unknown, we fit the data under $p = 0, \dots, P$, where P is a sufficiently larger number. The Akaike's information criterion (AIC) is used for model selection. The model with the smallest AIC value is considered as the best polynomial regression model for each process. Let p_i^* be the highest power term of the best model for process i , $i = 0, \dots, I$. The best estimated polynomial regression model for process i is

$$\hat{y}_{i,j} = \hat{\beta}_{i,0} + \hat{\beta}_{i,1} x_{i,j} + \cdots + \hat{\beta}_{i,p_i^*} (x_{i,j})^{p_i^*}, i = 0, \dots, I, j = 1, \dots, n \quad (3)$$

3.2 Test Statistics Calculation

To test the discrepancy between process g and process 0, the hypotheses are set as follows: the null hypothesis H_{0g} : process i has consistent performance as process 0, versus the alternative hypothesis H_{1g} : process i has inconsistent performance as process 0, $g = 1, \dots, I$. We propose the S_g test statistic that sums up the absolute standardized differences of the response variable at m evenly spaced values between a and b for the explanatory variable.

$$S_g = \sum_{j=1}^m \frac{|\hat{y}_{0,j} - \hat{y}_{g,j}|}{\sqrt{\hat{\sigma}_0^2 + \hat{\sigma}_g^2}}, g = 1, \dots, I \quad (4)$$

3.3 The Step-up Procedure

The I test statistics are arranged in an ascending order before applying the sequential tests. Let $S_{(1)} \leq S_{(2)} \leq \dots \leq S_{(I)}$ and $\{H_{0(g)}, H_{a(g)}\}$, $g = 1, \dots, I$, be the ordered statistics and their matching hypotheses, respectively. $H_{0(g)}$ is the null hypothesis that process (g) has consistent performance as process 0; $H_{a(g)}$ is the alternative hypothesis that process (g) has inconsistent performance as process 0, $g = 1, \dots, I$. For example, if $S_3 < S_1 < S_2$, processes (1), (2), and (3) refer to processes 3, 1, 2, respectively. The main concept of the step-up procedure is to test sequentially from the least significant hypothesis, and stop testing once rejecting a hypothesis. The test procedure starts with testing $\{H_{0(1)}, H_{a(1)}\}$, $H_{0(1)}$ is rejected if $S_{(1)}$ is greater than the critical value. Then, reject all of the untested hypotheses and stop the test procedure. Otherwise, accept $H_{0(1)}$ and continue testing $\{H_{0(2)}, H_{a(2)}\}$ until $\{H_{0(I)}, H_{a(I)}\}$ (Fig. 1). The processes being rejected are claimed

to have different quality performance from process 0. The processes being accepted are claimed to have consistent quality performance as process 0.

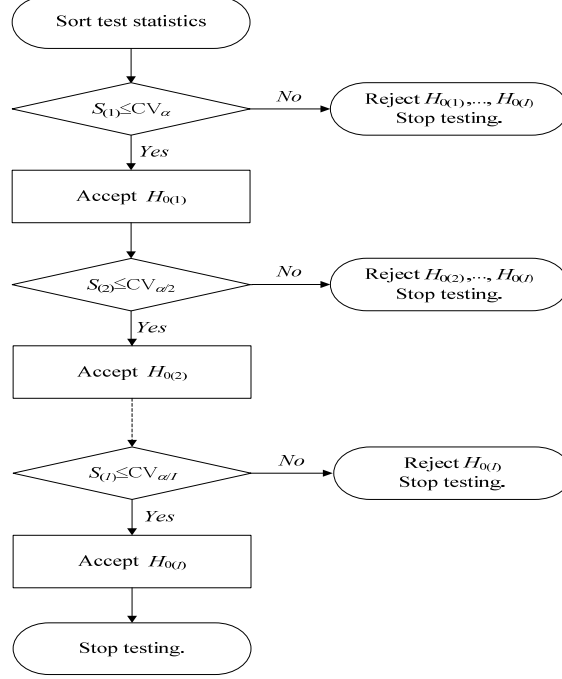


Fig. 1. The step-up test procedure.

The critical value, $(CV_{\alpha}, CV_{\alpha/2}, CV_{\alpha/I})$ is calculated by simulation. Two series of non-linear profile data from the estimated polynomial regression model of process 0 is generated under the null hypothesis. Then, calculate the test statistic S_g . Repeat for R times and sort R of the test statistics. The critical value, $CV_{\alpha/(g-1)}$, is the $\alpha/g * 100\%$ percentile of the sorted test statistics, $g = 1, \dots, I$. Based on the prove in [12], the step-up test scheme can control the overall error rate of no greater than α by modifying the significance level of each sequential test to α/g , $g = 1, \dots, I$.

4 Simulation Analysis

The power is the probability of rejecting the null hypotheses given that the alternative hypotheses are true. That is the probability of correctly rejecting the processes with inconsistent quality performance as process 0. To evaluate the performance of the proposed method, we analyze the power by simulation at $I = 2, 4, 6$. The underlying model of process 0 is set to $Y_{0,j} = 15 + 3X_{0,j} - 5(X_{0,j})^2 + 1.5(X_{0,j})^3 + \varepsilon_{i,j}$, $\varepsilon_{i,j} \sim N(0,1)$. Two shift patterns for alternative processes are investigated: (a) $E[Y_i|X] = (15 + i \delta_0) + 3X - 5X^2 + 1.5X^3$, (b) $E[Y_i|X] = 15 + (3X + i \delta_1) - 5X^2 + 1.5X^3$, $i = 1, \dots, I$ (Fig. 2). The value for the explanatory variable is within $a = 0$ and $b = 4$. The shift magnitude is set as δ_0

$= \delta_1 = 0.2, 0.4, \dots, 2$. $n = 10$ samples are collected for data fitting. The difference between profiles are calculated at $m = 10$ equally spaced x values between 0 and 4. The overall error rate is set to $\alpha = 0.05$. $R = 10^5$ simulation runs are performed under the null hypothesis to calculate the critical values (Table 1). A large number of simulation runs is applied to get stable critical values.

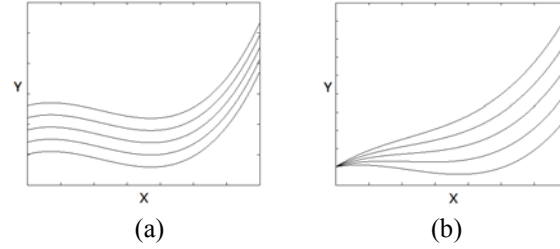


Fig. 2. The illustrations of the simulated mean shifts in (a) β_0 and (b) β_1 with $I = 4$.

Table 1. The critical values $CV_{0.05/g}$ for testing $I = 2, 4, 6$ processes, $g = 1, \dots, I$.

$I \setminus g$	1	2	3	4	5	6
2	5.1632	5.7967				
4	5.1632	5.7967	6.1763	6.4703		
6	5.1632	5.7967	6.1763	6.4703	6.6859	6.8511

The power is calculated as the average percentage of the processes being rejected among I different processes in 10^5 simulation runs. The power curves in Fig. 3 show that the power increases as the shift magnitude increases, which is reasonable. The power increases by I . This phenomenon is because that the simulation sets different nonlinear profile models for all of the alternative processes to be different from the existing process. The shift magnitude also increases by I . The analysis result suggests that the proposed method is effective in differentiating non-linear profiles for Process Selection.

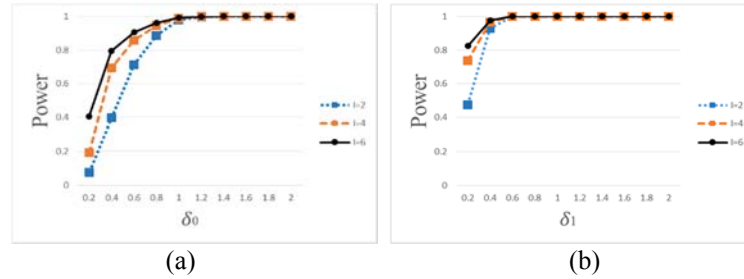


Fig. 3. The power curves under $I = 2, 4, 6$ with shifts in case (a) and (b).

5 Conclusion

Process Selection is a critical task in production management. Selecting and using multiple qualified alternative processes can increase production flexibility and capacity,

while maintaining quality. This research tackle the Process Selection problem by proposing the Step-up Test Procedure. By modifying the significance level of each sequential test along with using the step-up procedure, the proposed method controls the overall error rate to be no more than α . The proposed method effectively differentiate the non-linear profile of each alternative process from the existing process. Manufacturers can then use the selected processes along with the existing process to retain product quality in production. To relax the normality assumption, future research will investigate nonparametric methods for data fitting.

Acknowledgments

The authors would like to thank the reviewers for their valuable comments which greatly improve the quality of the paper.

References

1. Woodall, W. H., Montgomery, D. C.: Some current directions in the theory and application of statistical process monitoring. *Journal of Quality Technology*, 46(1), 78-94 (2014).
2. Woodall, W. H.: Bridging the gap between theory and practice in basic statistical process monitoring. *Quality Engineering*, 29(1), 2-15 (2017).
3. Gomaa, A. S., Birch, J. B.: A semiparametric nonlinear mixed model approach to phase I profile monitoring. *Communications in Statistics-Simulation and Computation*, 1-17 (2018).
4. Wang, F. K., Guo, Y.C.: Measuring process yield for nonlinear profiles. *Quality and Reliability Engineering International* 30(8), 1333-1339 (2014).
5. Chen, Y. J., Birch, J. B., Woodall, W. H.: A phase I cluster-based method for analyzing nonparametric profiles. *Quality and Reliability Engineering International* 31(8), 1675-1689 (2015).
6. Qiu, K., Wang, R., Peng, C., Lu, X., Wang, N.: Polynomial regression and interpolation of thermodynamic data in Al-Si-Mg-Fe system. *Calphad* 48, 175-183 (2015).
7. Xu, M., Pinson, P., Lu, Z., Qiao, Y., Min Y.: Adaptive robust polynomial regression for power curve modeling with application to wind power forecasting. *Wind Energy* 19(12), 2321-2336 (2016).
8. Fan, S. K., Chang, Y. J., Aidara, N. F.: Nonlinear profile monitoring of reflow process data based on the sum of sine functions. *Quality and Reliability Engineering International* 29(5), 743-758 (2013).
9. Fan, S.K., Lee, T.Y.: Using hyperbolic tangent function for nonlinear profile monitoring. In: *Proceedings of the Institute of Industrial Engineers Asian Conference*, pp. 457-463. Springer (2013).
10. Lee, J. S., Hur, Y. M., Kim, S. H., Wilson, J. R.: Monitoring nonlinear profiles using a wavelet-based distribution-free CUSUM chart. *International Journal of Production Research* 50(22), 6574-6594 (2012).
11. Wang, H. Z., Kim, S. H., Huo, X. M., Hur, Y. M., Wilson, J. R.: Monitoring nonlinear profiles adaptively with a wavelet-based distribution-free CUSUM chart. *International Journal of Production Research* 53(15), 4648-4667 (2015).
12. Hochberg, Y., Tamhane, A. C. *Multiple Comparison Procedures*, Wiley, New York (1987).