

Monitoring Simple Linear Profiles in the Leather Industry (A Case Study)

Amirhossein Amiri

**Department of Industrial Engineering
Shahed University, Tehran, Iran, amiri@shahed.ac.ir**

Ahmad Zand

**Department of Industrial Engineering
University of Science and Culture, Tehran, Iran, ahmad.zand66@gmail.com**

Darya Soudbakhsh

**Department of Industrial Engineering
Islamic Azad University-North Tehran Branch, Tehran, Iran, d.soudbakhsh@gmail.com**

Abstract

Sometimes, the quality of a process or product can be characterized by a relationship between a response variable and one explanatory variable, which is referred to as profile. In this paper, we introduce an example of a simple linear profile from the dyeing process of shoe leather in the leather industry and conduct a step-by-step Phase I analysis. For this purpose, we use two of the most powerful Phase I methods including the F method by Mahmoud and Woodall [1] and LRT method by Mahmoud et al. [2]. The results show that the investigated process is in statistical control.

Keywords

Color effluent, Dyeing process, Leather industry, Likelihood ratio test, Phase I

1. Introduction

In some situations, quality of a process or product is characterized by a relationship between a response variable and one or more explanatory variables which is referred to as profile. Many authors such as Kang and Albin [3], Mahmoud and Woodall [1], Mahmoud et al. [2], and Amiri et al. [4] presented applications of profile monitoring. We give here an example obtained from the leather industry. Leather is a popular material for making shoes. So, it is necessary to care about its quality. One of the processes which affects the quality of the leather is the dyeing process. The dyeing process is important in customer satisfaction because when the feet are in the shoes, the temperature of the shoes goes up and the feet begin to sweat. As a result, the color of the shoes begins to stain socks or feet. One of the most important quality characteristics in the dyeing process is the relationship between color effluent and temperature which should be monitored over time. The profiles which model this relationship are similar under in-control situation. On the other hand, out-of-control profiles indicate that the amounts of color effluent in one or more temperatures are not stable and there is assignable cause(s) in the dyeing process of shoe leather. The leather produced from the out-of-control dyeing process, could affect the shoe product under high temperatures and color effluent could lead to customer dissatisfaction (see Figure 1).

In the investigated case, no historical dataset was available. Hence, we dealt with Phase I analysis of profile monitoring. Our studies presented in the next sections show that the relationship between the color effluent of the leather and temperature can be well fitted by a simple linear regression model. The values of temperatures are selected based on the ones which could affect the final product. More explanations on data gathering are given in the second section.



Figure 1: The color effluent caused by leather shoe

Phase I analysis of profile monitoring is well studied by many authors. Mestek et al. [5], Stover and Brill [6], Kang and Albin [3], Mahmoud and Woodall [1], Mahmoud et al. [2] investigated Phase I monitoring of simple linear profiles. The purpose of Phase I analysis is to assess the stability of a process and to estimate the parameters. Mestek et al. [5] used a T^2 control chart in combination with principal component analysis (PCA) approach to monitor a simple linear profile in calibration application. Stover and Brill [6] proposed two methods for monitoring simple linear profiles. As a first method, they proposed a multivariate T^2 control chart. They proposed a PCA-based control scheme as a second method. Kang and Albin [3] investigated Phase I analysis of simple linear profiles by proposing two methods including T^2 and EWMA/R. Mahmoud and Woodall [1] suggested the use a global F-test to monitor the regression coefficients in conjunction with a univariate control chart for monitoring error standard deviation in Phase I. Mahmoud et al. [2] proposed a method to monitor linear profiles based on a likelihood ratio statistic. Zhu and Lin [7] proposed a shewhart-type control chart for monitoring slopes of linear profiles in both Phases I and II. The main goal in Phase II is to detect shift in the process parameters as quickly as possible. Authors including Kang and Albin [3], Kim et al. [8], Zou et al. [9], and Zhang et al. [10] studies Phase II monitoring of simple linear profiles. More complicated models such as multiple linear profiles, polynomial profiles, and nonlinear profiles are also investigated by some researchers such as Zou et al. [11], Mahmoud [12], Kazemzadeh et al. [13, 14], Williams et al. [15], Vaghefi et al. [16]. For survey papers in the area of profiles monitoring, authors are referred to Woodall et al. [17] and Woodall [18].

The structure of the paper is as follows: In the next section, data gathering is completely described. In the third section, it is shown the relationship can be modeled by a simple linear profile and model adequacy checking is conducted. Then, in the fourth section, Phase I studies are done using two of the most powerful methods in the literature, F method by Mahmoud and Woodall [1] and change point method by Mahmoud et al. [2], to investigate the stability of the process in Phase I. In the final section, our concluding remarks are given.

2. Data Gathering

As mentioned in the previous section, the performance of leather dyeing process is characterized by a relationship between the leather color effluent and temperature. To gather the data, the colored leathers, which are used to make leather shoes, were selected during 11 days and each one was divided to 5 pieces with the same size. These five pieces were examined in 150ml water at 5 different temperatures 25, 32, 39, 46 and 53°C and the corresponding color effluent for each temperature was measured. Hence, 11 profiles each one including 5 observations were gathered for Phase I analysis. To warm the water up to desired temperatures, we put it in the laboratory oven (See Figure 2). When the thermometer showed the specified temperatures, we dropped the leather into water and put the water in the oven again.

We took out the leather from water by a pincer after 2 hours. To measure the color effluent, we used a particular instrument called UV Spectrophotometer (shown in Figure 3). This instrument has two locations for liquids which one of them is used for the base liquid (in this case water) and the other one is for the liquid which is examined (in this case color effluent). To use the instrument, we filled the first place with water and the second one with color effluent selected each time from one of the 55 observations. Then the color effluent is measured and the data are saved in the instrument.

The result of these measurements was a dataset including 55 values which are summarized in Table 1.



Figure 2: Laboratory oven



Figure 3: UV Spectrophotometer

In this table, the first row shows the 5 values of the explanatory variable, temperature, and the first column shows the number of profiles (leathers). The values of response variable (color effluent) at 5 different temperatures are given through the table. These data were used for determining the model as well as Phase I monitoring in the following sections.

To recognize the model which is well-fitted to the data, first we plotted the scatter plot for all profiles (see Figure 4). The figure showed that a simple linear regression model is appropriate for the dataset. In addition, the scatter plot for each profile was plotted (not reported here) which showed the same result, simple linear regression model between color effluent and temperature.

Table 1: Dataset of the leather color effluent in 5 different temperatures

Temp.	25	32	39	46	53
Profile					
1	0.0218	0.02878	0.09083	0.10111	0.12566
2	0.0302	0.05422	0.07183	0.11716	0.13127
3	0.0288	0.02868	0.08575	0.09310	0.13549
4	0.0306	0.07571	0.01011	0.11624	0.12850
5	0.0488	0.02806	0.08549	0.11812	0.11880
6	0.0310	0.09438	0.07157	0.11922	0.14965
7	0.0231	0.07626	0.08093	0.13988	0.15714
8	0.0455	0.09253	0.15109	0.08746	0.14101
9	0.0209	0.04746	0.10231	0.12651	0.12299
10	0.0578	0.02227	0.11557	0.11261	0.09202
11	0.0463	0.06435	0.08679	0.07877	0.10632

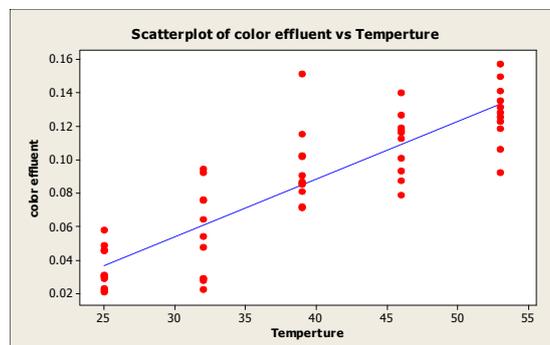


Figure 4: Scatter plot for all profiles

3. Model and structure of the process

After specifying the model, simple linear regression model fitted to each profile. The estimators of regression parameters, analysis of variance table, coefficient of determination, adjusted coefficient of determination, variance inflation factor (VIF) and the results of lack of fit test for the first profile are reported in Table 2. The results show that the adjusted coefficient of determination is equal to 90.3% and the model is well fitted. The Statistical hypothesis test also confirms this issue. In analysis of variance table as well as separate individual tests, the significance of the regression parameters is confirmed. There is no evidence of lack-of-fit since the p-value of the test is greater than 0.05. It should be noted that we used the lack-of-fit test proposed first by Burn and Ryan [19] due to no replication in explanatory variables. The same results (not reported here) were obtained for the other profiles. The normal probability plot for standardized residuals of each profile (not reported here) also confirmed the normality assumption of the error terms.

Table 2: Regression analysis: color effluent versus temperature

Regression Analysis: color effluent 1 versus temperature					
The regression equation is					
color effluent 1 = - 0.0823 + 0.00400 temperature					
Predictor	Coef	SE Coef	T	P	VIF
Constant	-0.08228	0.02601	-3.16	0.051	
Temperature	0.0039984	0.0006463	6.19	0.009	1.000
S = 0.0143070 R-Sq = 92.7% R-Sq(adj) = 90.3%					
Analysis of Variance					
Source	DF	SS	MS	F	P
Regression	1	.0078337	0.0078337	38.27	0.009
Residual Error	3	0.0006141	0.0002047		
Total	4	0.0084478			
No evidence of lack of fit (P >= 0.1).					

4. Phase I Analysis

As mentioned in the introduction section, there are many Phase I methods for monitoring simple linear profiles in the literature. In this paper, F and LRT methods proposed by Mahmoud and Woodall [1] and Mahmoud et al. [2], respectively, are used for Phase I analysis because these methods are more powerful than the competing methods in the literature to detect step shifts and outliers in Phase I. In this section, we first explain about the methods and then apply the methods on the available dataset shown in Table 1.

4.1 F method

In this method which is first proposed by Mahmoud and Woodall [1], one first tests for the equality of regression lines. To do so, one pools all the m samples into one sample of size N and creates $m-1$ indicator variables such that

$$\begin{cases} Z_{ji} = 1 & \text{if observation } i \text{ is from sample } j \\ Z_{ji} = 0 & \text{otherwise, } i = 1, 2, \dots, N \quad j = 1, 2, \dots, m-1 \end{cases}$$

then, one fits the following simple linear regression to the pooled data:

$$y_i = \beta_0 + \beta_1 X_{1i} + \beta_{01} Z_{1i} + \beta_{11} Z_{1i} X_{1i} + \dots + \beta_{0m'} Z_{m'i} + \beta_{1m'} Z_{m'i} X_{1i} + \varepsilon_i \quad i = 1, 2, \dots, N \quad m' = m-1 \quad (1)$$

To test for equality of the m regression lines, we test the hypotheses $H_0: \beta_{01} = \beta_{02} = \dots = \beta_{0m'} = \beta_{11} = \beta_{12} = \dots = \beta_{1m'} = 0$ versus $H_1: H_0$ is not true.

Under the null hypothesis, the reduced model is as follows:

$$y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \quad i = 1, 2, \dots, N \quad (2)$$

The standard test statistic for testing H_0 is

$$F = \frac{[SSE(R) - SSE(F)] / 2(m-1)}{SSE(F) / (N - 2m)} \quad (3)$$

where $SSE(F)$ and $SSE(R)$ are the residual sum of squares resulting by fitting the regression models in Equations (1) and (2), respectively. The statistic in Equation (3) has an F distribution with $2(m-1)$ and $N-2m$ degrees of freedom under the null hypothesis.

Mahmoud and Woodall [1] proposed the use of a univariate control chart in conjunction with Global F test in Equation (3) to monitor standard deviation of simple linear profiles. They assumed that the sample sizes are constant from sample to sample as it is in our case. The statistics which are plotted on the univariate control chart are:

$$F_j = \frac{MSE_j}{\sum_{i \neq j}^m MSE_i / (m-1)} \quad (4)$$

The statistics in the above equation follow an F distribution with $(n-2)$ and $(m-1)(n-2)$ degrees of freedom under the null hypothesis. The lower and upper control limits for the F statistics are as follows:

$$LCL = F_{(n-2), (n-2)(m-1), 1-\alpha/2} \quad \text{and} \quad UCL = F_{(n-2), (n-2)(m-1), \alpha/2} \quad (5)$$

We perform the global F test at significance level of $\alpha_1 = 0.0202041$ and set the control limits of the univariate control chart based on a false-alarm probability of $\alpha_2 = 0.00185$ to obtain overall probability of Type I error equals to 0.04. The mentioned probabilities are computed as follows:

$$\alpha_1 = 1 - (1 - \alpha_{\text{overall}})^{1/2} = 1 - (1 - 0.04)^{1/2} = 0.0202041$$

$$\alpha_2 = 1 - (1 - \alpha_1)^{1/m} = 1 - (1 - 0.0202041)^{1/11} = 0.00185$$

In this case, the F statistics for testing the equality of all the regression lines is equal to 0.9535 and has F distribution with 20 and 33 degrees of freedom. Hence, at the significance level of α_1 equals to 0.0202041, $F_{20,33,0.0202041}$ is 2.22547 and there is no reason to reject the null hypothesis; i.e. all of the regression lines are identical.

Also, the control limits for testing the stability of the error term variance are $LCL=0.0075656$ and $UCL=7.14710$. As shown in Figure 5, all of the sample statistics obtained by Equation (4) are within the control limits. Therefore, we can conclude that the process variance is stable.

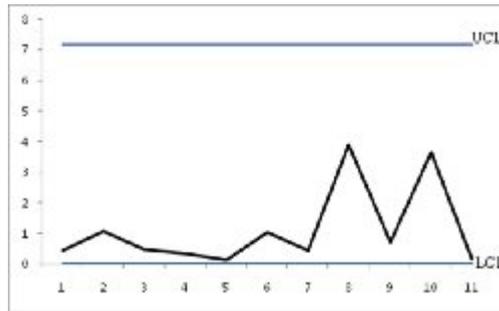


Figure 5: Control chart for the error term variance

4.2. Likelihood ratio test method

The second method is based on the likelihood ratio test first proposed by Mahmoud et al. [2] for monitoring simple linear profiles in Phase I. With a step shift in one or more regression parameters after sample m_1 , the following models are assumed:

$$Y_{ij} = \beta_{01} + \beta_{11}x_{ij} + \varepsilon_{ij1} \quad i=1,2,\dots,n \quad j=1,2,\dots,m_1 \quad \text{and} \quad Y_{ij} = \beta_{02} + \beta_{12}x_{ij} + \varepsilon_{ij2} \quad i=1,2,\dots,n \quad j=m_1+1,\dots,m, \quad (6)$$

where $\varepsilon_{ij1} \sim N(0, \sigma_1^2)$ and $\varepsilon_{ij2} \sim N(0, \sigma_2^2)$.

To test the following hypotheses

$$H_0 : \beta_{01} = \beta_{02} = \beta_0, \quad \beta_{11} = \beta_{12} = \beta_1, \quad \text{and} \quad \sigma_1^2 = \sigma_2^2 = \sigma^2 \quad H_1 : H_0 \text{ is not true.}$$

One first obtains the following likelihood ratio statistic, lrt_{m_1} , for all possible values of $m_1 = 1, 2, \dots, m-1$:

$$lrt_{m_1} = N \log \hat{\sigma}^2 - N_1 \log \hat{\sigma}_1^2 - N_2 \log \hat{\sigma}_2^2, \quad (7)$$

where $\hat{\sigma}^2$ is maximum likelihood estimator of the error variance for the simple linear regression model fitted to all the m samples pooled into one sample of size N and computed as follows:

$$\hat{\sigma}^2 = \sum_{i=1}^N (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2 / N. \quad (8)$$

$\hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$ are the maximum likelihood estimators of the error variances for the simple linear regression fitted to all the samples prior to and following m_1 pooled into one sample of size $N_1 = \sum_{j=1}^{m_1} n_j$ and $N_2 = N - N_1$, respectively and given by

$$\hat{\sigma}_1^2 = \sum_{i=1}^{N_1} (y_i - \hat{\beta}_{01} - \hat{\beta}_{11} x_i)^2 / N_1 \text{ and } \hat{\sigma}_2^2 = \sum_{i=N_1+1}^N (y_i - \hat{\beta}_{02} - \hat{\beta}_{12} x_i)^2 / (N - N_1). \quad (9)$$

In Equations (7) to (9), the regression parameters are estimated by using ordinary least square method. Then one divides each likelihood ratio statistic by its expected value under the null hypothesis which can be approximated using the following equation:

$$E(lrt_{m1}) \approx 2 - 2\left(\frac{1}{N} - \frac{1}{N_1} - \frac{1}{N_2}\right) - \left(\frac{N}{N-2} - \frac{N_1}{N_1-2} - \frac{N_2}{N_2-2}\right) - \frac{1}{3} \left(\frac{N}{(N-2)^2} - \frac{N_1}{(N_1-2)^2} - \frac{N_2}{(N_2-2)^2}\right), \quad (10)$$

and obtains the following statistics:

$$lrtc_{m1} = lrt_{m1} / E(lrt_{m1}). \quad (11)$$

The method signals the presence of one or more assignable causes if the maximum of the statistics in Equation (11) exceeds a threshold. Mahmoud et al. [2] proposed an approximation for the threshold to give a specified probability of Type I error. Based on Mahmoud et al. [2], the threshold is approximated by $\chi_{3,1-(\alpha/r^*)}^2 / 3$ for $m > 6$ where

$\chi_{3,1-(\alpha/r^*)}^2$ is the $100(1-(\alpha/r^*))$ percentile of the chi-squared distribution with three degrees of freedom and r^* is:

$$r^*(m) = -11.5 + 8.05 \log m. \quad (12)$$

The likelihood ratio test statistics, their corresponding expected values, and corrected likelihood ratio statistics are computed by Equations (7), (10), and (11), respectively and the results are summarized in Table 3.

To calculate the threshold, we first should compute r^* by using Equation (12) which is equal to 7.8 in the considered case. Then, the threshold would be $\chi_{3,1-(\alpha/r^*)}^2 / 3$ which is equal to 4.2613. Since the $\max(lrtc_{m1})$, which is equal to 3.5399 (see Table 3), does not exceed the threshold, 4.2613, it can be concluded that the process is stable and in statistical control.

Table 3: lrt_{m1} , $lrtc_{m1}$ statistics and the expected value of lrt_{m1}

m_1	lrt_{m1}	$E(lrt_{m1})$	$lrtc_{m1}$
1	4.5992	4.2601	1.0796
2	3.6544	3.5205	1.038
3	7.7703	3.348	2.3209
4	6.5059	3.2795	1.9838
5	11.5152	3.2529	3.5399
6	9.1678	3.2529	2.8138
7	10.281	3.2795	3.1349
8	3.6763	3.348	1.0981
9	7.1038	3.5205	2.0178
10	10.0253	4.2601	2.3533

Considering the fact that both F and LRT methods showed that the process is stable and in-control, the dataset can be used to estimate the regression parameters and control charts for Phase II monitoring can be constructed.

5. Conclusion

In this paper, we investigated the relationship between color effluent and temperature of the dying process from the leather industry. Our analyses showed that this relationship is a simple linear regression and adequacy of the model is satisfied. Two of the most powerful methods in the literature of profiles monitoring are applied for Phase I

analyses. The results showed that the process is in-control and the dataset can be used to estimate the parameters and control limits for monitoring of the process in Phase II. Designing a method for monitoring this relationship in Phase II can be considered as a future research.

References

1. Mahmoud, M.A., and Woodall, W.H., 2004, "Phase I Analysis of Linear Profiles With Calibration Applications," *Technometrics*, 46(4), 380-391.
2. Mahmoud, M.A., Parker, P.A., and Woodall, W.H., 2007, "A Change Point Method for Linear Profile Data," *Quality and Reliability Engineering International*, 23(2), 247-268.
3. Kang, L., and Albin, S.L., 2000, "On-line Monitoring When the Process Yields a Linear Profile," *Journal of Quality Technology*, 32(4), 418-426.
4. Amiri, A., Jensen, W.A., and Kazemzadeh, R.B., 2010, "A Case Study on Monitoring Polynomial Profiles in the Automotive Industry," *Quality and Reliability Engineering International*, 26(5), 509-520.
5. Mestek, O., Pavlik, J., and Suchánek, M., 1994, "Multivariate Control Charts: Control Charts for Calibration Curves," *Fresenius, Journal of Analytical Chemistry*, 350(6), 344-351.
6. Stover, F.S., and Brill, R.V., 1998, "Statistical Quality Control Applied to Ion Chromatography Calibrations," *Journal of Chromatography A*, 804(1-2), 37-43.
7. Zhu, J., and Lin, D.K.J., 2010, "Monitoring the Slopes of Linear Profiles," *Quality Engineering*, 22(1), 1-12.
8. Kim, K., Mahmoud, M.A., and Woodall, W.H., 2003, "On the Monitoring of Linear Profiles," *Journal of Quality Technology*, 35(3), 317-328.
9. Zou, C., Zhang, Y., and Wang, Z., 2006, "Control Chart Based on Change-Point Model for Monitoring Linear Profiles," *IIE Transactions*, 38(12), 1093-1103.
10. Zhang, J., Li, Z., and Wang, Z., 2009, "Control Chart Based on Likelihood Ratio for Monitoring Linear Profiles," *Computational Statistics and Data Analysis*, 53(4), 1440-1448.
11. Zou, C., Tsung, F., and Wang, Z., 2007, "Monitoring General Linear Profiles Using Multivariate Exponentially Weighted Moving Average Schemes," *Technometrics*, 49(4), 395-408.
12. Mahmoud, M.A., 2008, "Phase I Analysis of Multiple Linear Regression Profiles," *Communications in Statistics, Simulation and Computation*, 37(10), 2106-2130.
13. Kazemzadeh, R.B., Noorossana, R., and Amiri, A., 2008, "Phase I Monitoring of Polynomial Profiles," *Communications in Statistics-Theory and Methods*, 37(10), 1671-1686.
14. Kazemzadeh, R.B., Noorossana, R., and Amiri, A., 2009, "Monitoring Polynomial Profiles in Quality Control Applications," *The International Journal of Advanced Manufacturing Technology*, 42(7), 703-712.
15. Williams, J. D., Woodall, W.H., and Birch, J.B., 2007, "Statistical Monitoring of Nonlinear Product and Process Quality Profiles," *Quality and Reliability Engineering International*, 23(8), 925-941.
16. Vaghefi, A., Tajbakhsh, S.D., and Noorossana, R., 2009, "Phase II Monitoring of Nonlinear Profiles," *Communications in Statistics-Theory and Methods*, 38(11), 1834-1851.
17. Woodall, W.H., Spitzner, D.J., Montgomery, D.C., and Gupta, S., 2004, "Using Control Charts to Monitor Process and Product Quality Profiles," *Journal of Quality Technology*, 36(3), 309-320.
18. Woodall, W.H., 2007, "Current Research on Profile Monitoring," *Revista Produção*, 17(3), 420-425.
19. Burn, D.A., and Ryan, T.A.Jr., 1983, "A diagnostic test for lack of fit in regression models," *ASA Proc. of the Statistical Computing Section*, 286-290.