Robust Phase I Monitoring of Profile Data with Application in Low-E Glass Manufacturing Processes

Li Zeng[a*†], Smriti Neogi[a] and Qiang Zhou[b]

[a]*Department of Industrial and Manufacturing Systems Engineering, The University of Texas at Arlington, Arlington, TX, 19017, U.S.A.*
[b]*Department of Systems Engineering and Engineering Management, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon, Hong Kong Special Administrative Region*
[*]*Correspondence to: Li Zeng, Department of Industrial and Manufacturing Systems Engineering, The University of Texas at Arlington, 500 West First Street, PO Box 19017, Arlington, TX 76019, U.S.A.*
[†]*E-Mail: lzeng@uta.edu*

Normality is usually assumed in profile monitoring. However, there are many cases in practice where normality does not hold. In such cases, conventional monitoring techniques may not perform well. In this study, we propose a robust strategy for Phase I monitoring of quality profile data in the presence of non-normality. This strategy consists of three components: modeling of profiles, independent component analysis (ICA) to transform multivariate coefficient estimates in profile modeling to independent univariate data, and univariate nonparametric control charts to detect location/scale shifts in the data. Two methods for multiple change point detection are also studied. The properties of the proposed method are examined in a numerical study and it is applied to optical profiles from low-E glass manufacturing in the case study.

Keywords: profile monitoring; Phase I monitoring; independent component analysis (ICA); nonparametric control charts; non-normality

1. Introduction

In some manufacturing processes, product quality is characterized by the relationship between a response variable and an explanatory variable, called *profiles*. Monitoring of quality profiles has received much attention recently due to the increasing popularity of this type of quality data in practice[1]. Parametric and nonparametric methods have been developed for this purpose. This study focuses on parametric profile monitoring where the shape of the profiles can be characterized by a parametric model adequately.

The basic idea of parametric profile monitoring includes two steps: First, an appropriate statistical model is used to characterize the profiles. The choice of models depends on the

characteristics of the profile data in the studied applications. Linear models[2], polynomial models[3-6], splines[7], mixed-effect models[8-9] and nonlinear models[10] have been used in existing studies. Second, the parameter estimates of the fitted model are monitored by using multivariate control chart techniques such as T^2 control chart and Multivariate EWMA control charts[11]. Woodall[12] and Noorossana *et al.*[11] give excellent review of the state of art in this research area.

The majority of existing studies on profile monitoring focus on Phase II monitoring, while only a few efforts are made on Phase I analysis. Mahmoud and Woodall[13] develop an *F* test approach for Phase I monitoring of linear profiles. Mahmoud *et al.*[14] propose a change point method based on likelihood ratio test for Phase I monitoring of linear profiles. Kazemzadeh *et al.*[15] compare three approaches for Phase I analysis of polynomial profiles, including the extension of the change point method, the extension of the *F* test approach, and a standard procedure based on T^2 test. It is found that the change point approach performs the best.

One limitation of the literature is that most studies rest on normality assumptions: the random errors in the profile models are typically assumed to be normally distributed, and the random effects in mixed-effect models are also bound by this assumption. However, this may not be the case in some manufacturing processes, such as the low-emittance (low-E) glass manufacturing process illustrated in Figure 1. The low-E glass is a type of energy-efficient glass products which is manufactured through physical or chemical coating processes, where solid materials, e.g., metal, metal oxide and metal nitride, are deposited on the surface of flat glass. The coating enhances the thermal/optical performance of the product so that they are able to reduce unwanted heat gain in summer and heat loss in winter[16-18]. The quality of coating is measured by optical profiles of scanned locations on the glass surface. Figure 1 shows an example of a typical type of optical profiles, the reflectance profiles, which represent the percentage of light (*r*) that reflects from the glass surface over a range of wavelengths (*λ*). Due to the many chemical subprocesses involved in low-E glass manufacturing, various random noises may be present in the production. As a result, the quality measurements may contain a considerable amount of extreme values and thus normality assumptions are not appropriate in modeling such profiles.

In fact, the effect of non-normality on the performance of profile monitoring has been investigated by a group of researchers such as Mahmoud and Woodall[13], Williams *et al.*[19], Vaghefi *et al.*[20] and Noorossana *et al.*[21] A general conclusion in these studies is that when normality is not satisfied, the conventional profile monitoring techniques may give misleading

results. Though some techniques are found to be robust for certain types of deviations from normality[21], there lacks a generic method for profile monitoring in the presence of non-normality.
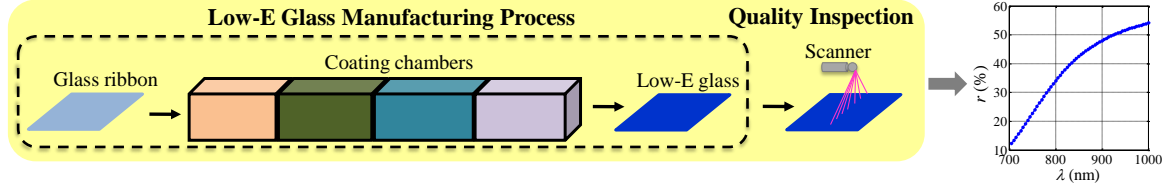

Figure 1. Low-E glass manufacturing process and example of optical profiles

This study aims to fill the gap in the literature by proposing a robust strategy for Phase I analysis of quality profiles. For non-normal data, nonparametric control charts are usually used to replace the conventional control charts based on normality. This idea is adopted in the proposed strategy. Moreover, to avoid issues with multivariate monitoring, independent component analysis (ICA) is used to transform multivariate coefficient estimates of the profile models into univariate independent data. In addition, we also study two methods to detect multiple change points as this is often the case in Phase I analysis. The properties of this strategy are demonstrated in a numerical study considering different scenarios of non-normality. In the case study, it is applied to optical profile data from low-E glass manufacturing as shown in Figure 1.

The remainder of the paper is organized as follows. Section 2 presents the problem formulation in Phase I analysis of profiles and the basic idea of the proposed strategy. Details of each component in the strategy are given in Section 3. Section 4 and 5 report the results of the numerical study and the case study respectively. Finally, Section 6 summarizes the findings in this work.

2. Problem Formulation

2.1 Profile models and Phase I monitoring

Without loss of generality, polynomial models which fit the optical profiles in Figure 1 will be used to illustrate the proposed method. Let x be the explanatory variable and y be the corresponding response. Suppose there are m profiles, each containing n sampling points, and the x values are fixed and constant among the profiles. Two types of polynomial models have been used in the literature: regular polynomial models[3-5] and mixed-effect polynomial models[6]. Mathematical expressions of these models are given below:

Regular polynomial model

$$y_{ij} = \beta_p x_j^p + ... + \beta_h x_j^h + ... + \beta_0 + \varepsilon_{ij} \qquad (1)$$

where $i=1,...,m$ is the index of profiles, $j=1,...,n$ is the index of sampling points, and $h=0,...,p$ is the index of the exponent of polynomials. $\beta_0,..., \beta_p$ are the fixed, unknown coefficients, and ε_{ij} is the random error which follows certain non-normal distribution with zero mean.

Mixed-effect polynomial model

$$y_{ij} = C_p x_j^p + ... + C_h x_j^h + ... + C_0 + \varepsilon_{ij}$$
$$C_h = \eta_h + \alpha_{hi} \qquad (2)$$

where each coefficient C_h consists of two parts: the *fixed effect*, η_h, which is fixed and unknown, and the random effect, α_{hi}, which varies from profile to profile. The random effects are assumed to follow non-normal distributions with zero mean. The mixed-effect models are preferred when the within-profile correlation is significant[6,8] or when there is intrinsic variation in the shapes of profiles.

We take a change-point view in the Phase I analysis, that is, assume the historical data stream contains w change points, i.e.,

$$
\begin{aligned}
y_{ij} &\sim M_1 & &\text{for } 1 \le i \le K_1; \\
y_{ij} &\sim M_2 & &\text{for } K_1 + 1 \le i \le K_2; \\
&\vdots & &\qquad \vdots \\
y_{ij} &\sim M_{w+1} & &\text{for } K_w + 1 \le i \le m.
\end{aligned}
\qquad (3)
$$

where $K_1,..., K_w$ are the change points, and $M_1,..., M_{w+1}$ are the polynomial models followed by the data between adjacent change points. Two types of changes may occur in the data: *location shift* which corresponds to the change in the coefficients in (1) or the fixed effects in (2), and *scale shift* which corresponds to the change in the scale of the random error in (1) or the scale of random effects/error in (2). The goal of Phase I monitoring is three-fold: determine whether any change occurs in the data, identify the change points as accurately as possible, and establish in-control parameters based on the change point estimates.

2.2 Basic idea of the proposed strategy

Following the standard practice of parametric profile monitoring, change detection will be applied to the coefficient estimates of the profile models, i.e., βs in (1) or Cs in (2). When those estimates are not normally distributed, a natural idea is to use nonparametric control charts. As multiple coefficients typically exist, we can either use a multivariate nonparametric control chart on all the coefficients simultaneously, or use univariate nonparametric control charts on each coefficient separately. Since the estimates of coefficients are correlated with each other, the first

4

solution appears to be more reasonable. It is also possible since multivariate nonparametric control charts are available in the literature, including the sign MEWMA chart proposed by Zou and Tsung[22], and the rank-based MCUSUM chart and the nonparametric MCUSUM chart proposed by Qiu and Hawkins[23,24]. However, as calculating the statistics in these multivariate techniques involves matrix inversion operations, they may suffer instability issues in some cases. For example, according to our simulations, when the variation of the coefficient estimates is not balanced, that is, the variation of some coefficients is too large or too small, the statistics may not exist due to singularity of inverted matrices. In addition, the results from multivariate control charts do not have easy interpretation. In contrast, the second idea is free of instability issues and easy to implement and understand, given that the correlation between the estimates of coefficients can be eliminated in some way.

In fact, the second idea is followed in the study of Kazemzadeh *et al.*[4] for Phase II monitoring of polynomial profiles under normality assumptions, where orthogonal polynomial models are used for fitting the profile data. Since the coefficient estimates in orthogonal polynomial regression are independent, they can be monitored separately using univariate control charts. In this study, we propose a similar method to solve this problem in the context of non-normality, which uses independent component analysis (ICA) to transform the multivariate coefficient estimates into univariate independent components (ICs), and then applies univariate nonparametric control charts to each IC. This method is generic in that it can be applied to different forms of polynomial models and other models. Moreover, as will be explained in Section 3.2, the use of ICA will bring special benefits in change point detection.

Figure 2 shows the components of the proposed strategy for Phase I monitoring of profile data. First, polynomial models are fitted for the data to obtain the estimates of coefficients. Second, ICA is applied to the multivariate coefficient estimates. The selected ICs are then monitored using univariate nonparametric control charts to detect location/scale shifts. Considering that multiple change points may exist, once a change point is detected, the data stream will be segmented at the change point and the detection will continue on the uninspected data. Details of each component will be given in Section 3.
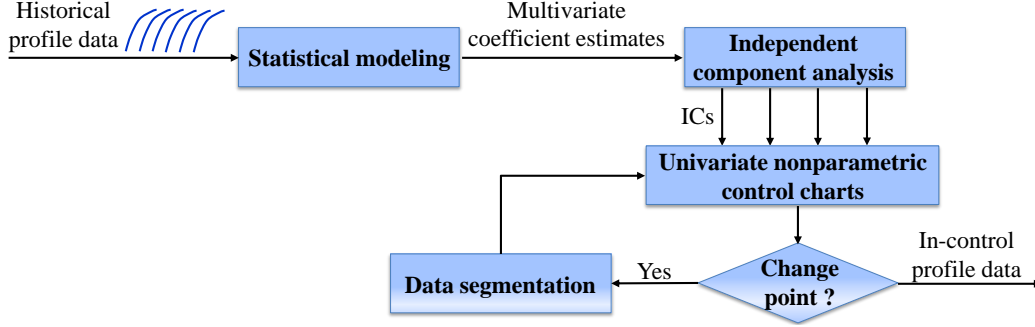
5

Historical profile data → **Statistical modeling** → Multivariate coefficient estimates → **Independent component analysis**

ICs → **Univariate nonparametric control charts**

Data segmentation ← Yes ← **Change point ?** → In-control profile data

Figure 2. The basic idea of the proposed strategy for Phase I monitoring of profile data

3. The Proposed Strategy for Phase I Monitoring of Profile Data

3.1 Statistical modeling

The first step in the Phase I analysis is to fit a polynomial model for each profile in the historical data stream. The degree of polynomials, p, can be determined through preliminary analysis comparing the residuals under different choices of p. Since non-normality is assumed, ordinary least squares method can be used to fit the models. If the underlying model is a regular polynomial model in (1), the estimates of coefficients are

$$\begin{bmatrix} \hat{\beta}_{p1} & \cdots & \hat{\beta}_{01} \\ \vdots & & \vdots \\ \hat{\beta}_{pm} & \cdots & \hat{\beta}_{0m} \end{bmatrix}$$

where each column represents the estimates of one coefficient from the m profiles. For mixed-effect models in (2), the obtained matrix represents the estimates of Cs. The coefficient estimates will be used in the following analyses.

3.2 Independent component analysis

ICA is a data projection technique which transforms original multivariate data into univariate independent components through linear transformation[25]. Similar to another popular projection tool, the principle component analysis (PCA), ICA is often used for dimension reduction purposes in the literature as a number of significant ICs can be selected to represent the original data. For example, it is used in monitoring complex nonlinear profiles to reduce the dimension of data points contained in each profile[26]. Ding *et al.*[26] point out another advantageous aspect of ICA: unlike PCA which projects data onto a lower subspace that preserves the majority of the variability in the original data, ICA projects data to a subspace where the distinction of any existing structures in the data will be maximized in the resulting ICs. So the objective of ICA

aligns well with the objective of Phase I analysis, i.e., separating data following different structures.

With the abovementioned properties, ICA is appropriate in our study to transform the multivariate coefficient estimates into univariate independent components, so that univariate nonparametric control charts can be applied to detect changes. Some algorithms of ICA are available in commercial software such as Matlab and R. The *fastICA* function in Matlab is used in this study. It is worth mentioning that as the degree of polynomials is typically not high, the advantage of ICA in data reduction is not a key concern here. Results of the numerical study show clearly its role in manifesting the changes in the data, as given in Section 4.

3.3 Univariate nonparametric control chart

Various univariate nonparametric methods have been developed for monitoring non-normal data, including the bootstrap control chart by Jones and Woodall[27], and the rank-based tests by Gordon and Pollak[28], and Hackl and Ledolter[29]. Here we choose the control chart proposed by Hawkins and Deng[30] for detecting location shifts and the one proposed by Ross *et al.*[31] for detecting scale shifts. These two techniques are chosen because they do not require prior knowledge of in-control parameters and easy to implement. Moreover, they can also be applied for Phase II monitoring of large-volume data streams which exist in many manufacturing processes such as the low-E glass manufacturing. Note that these techniques are designed to detect a single change point in the data; detection of multiple change points is realized through data segmentation which will be described in Section 3.4. The basics of the two techniques are provided as follows.

Assume $Z_1,\ldots, Z_m$ are independent non-normal random variables with distribution

$$\begin{aligned} Z_i &\sim F_1 \qquad \text{for } 1 \leq i \leq K; \\ Z_i &\sim F_2 \qquad \text{for } K+1 \leq i \leq m. \end{aligned}$$

where K is the change point between the two different distributions F_1 and F_2. The focus here is to determine whether a change exists and if so, estimate the change point K.

The control chart of Hawkins and Deng[30] to detect location shifts is based on the Mann-Whitney two-sample test. Let

$$D_{ij} = \text{sgn}(Z_i - Z_j) = \begin{cases} 1 & \text{if } Z_i > Z_j \\ 0 & \text{if } Z_i = Z_j \\ -1 & \text{if } Z_i < Z_j \end{cases}$$

7

where $1 \leq i, j \leq m$. The Mann-Whitney statistic is defined based on the D_{ij},

$$U_{k,m} = \sum_{i=1}^{k} \sum_{j=k+1}^{m} D_{ij}$$

for $1 \leq k \leq m-1$. The standardized version of this statistic is

$$U'_{k,m} = \frac{U_{k,m}}{\sqrt{k(m-k)(m+1)/3}}$$

which follows a standard normal distribution asymptotically. Note that this statistic holds for each value of k. A natural estimate of the change point is the value of k that gives the largest $U'_{k,m}$. In the control chart, the following statistic is used

$$\begin{aligned} U'_{\max,m} &= \max_{1 \leq k \leq m-1} \left| U'_{k,m} \right| \\ \hat{K} &= \arg \max_{1 \leq k \leq m-1} \left| U'_{k,m} \right| \end{aligned} \tag{4}$$

The control limit needs to be found through simulations under a specified in-control average run length ($\text{ARL}_{\text{in-control}}$). It is required that $m \geq 15$.

The control chart of Ross *et al.*[31] to detect scale shifts is based on the Mood test. The Mood statistic is

$$M_{k,m} = \sum_{i=1}^{k} \left(R_i - \frac{m+1}{2} \right)^2$$

where R_i is the rank of Z_i among $\{Z_1, \ldots, Z_k\}$. The standardized version is

$$M'_{k,m} = \frac{\left| M_{k,m} - k(m^2-1)/12 \right|}{\sqrt{k(m-k)(m+1)(m^2-4)/180}}$$

The statistic of the control chart takes a similar form as the Mann-Whitney statistic in (4),

$$\begin{aligned} M'_{\max,m} &= \max_{1 \leq k \leq m-1} \left| M'_{k,m} \right| \\ \hat{K} &= \arg \max_{1 \leq k \leq m-1} \left| M'_{k,m} \right| \end{aligned} \tag{5}$$

The control limit also needs to be obtained through simulations. Fortunately, Ross *et al.*[31] provides polynomial approximations of the control limit under a group of in-control average run lengths. It is required that $m \geq 20$.

3.4 Data segmentation for multiple change point detection

To identify multiple change points that may exist in the data, the two control charts in Section 3.3 need to be used repeatedly. There are two ways to do this as illustrated in Figure 3:

- *Binary segmentation* (BS): Change detection is first conducted on all the data. Whenever a change is detected, the data stream is split into two segments at the estimated change point. Then change detection is conducted on each segment separately.

- *Sequential segmentation* (SS): Change detection is conducted sequentially starting from the segment with minimum required number of data points. If no change is detected, a new data point will be added to the segment and the detection continues; when a change is detected, the segment by the estimated change point will be discarded and change detection is applied to the subsequent data.
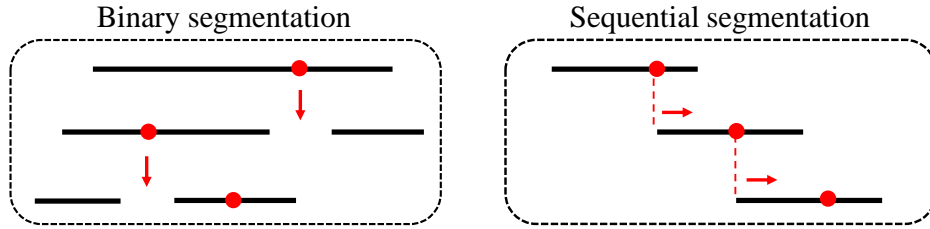

Figure 3. Two ways for data segmentation in multiple change point detection

Each of the two methods has been used in existing studies for detecting multiple change points[3,31], but no study has been done to evaluate and compare their performance. In general, they both have pros and cons: the BS method works on a whole segment to detect changes, while the SS method adds new data point one by one. So the sample size in the BS method is likely to be larger than in the SS method, and thus the BS method tends to be more accurate in identifying the change points; on the other hand, the segment used in the BS method may contain multiple change points, while that used in the SS method is more likely to contain one single change point due to its sequential nature. So the assumption of single change point holds better for the SS method, and thus it is supposed to be more accurate. Simulation results on the performance of these two methods will be given in Section 4.2.

4. Numerical Study

Simulations are done to address the following concerns:

1. *Performance of the two data segmentation methods described in Section 3.4 in multiple change point detection, and*

2. *Properties of the proposed strategy for Phase I monitoring of profile data.*

For the first concern, univariate non-normal data streams containing two change points are simulated under different parameter scenarios, and the two data segmentation methods are

applied to each stream. Their performance in identifying the true change points is evaluated and compared. For the second concern, profile data with non-normal errors are simulated under different parameter scenarios, and the proposed Phase I analysis is applied. Characteristics of the proposed strategy will be summarized. In this section, we will first describe how data are generated in the simulations, and then report the results of the above studies.

4.1 Data generation

Univariate data following non-normal distributions need to be simulated in this study. To be flexible, we use two large classes of non-normal distributions, the skew-normal distribution[32] and the skew-t distribution[33], which represent general cases of skewed and/or heavy-tailed distributions. For a random variable Z following the skew-normal distribution $SN(\mu, \sigma^2, \lambda)$ with location parameter μ, scale parameter σ^2 and skewness parameter λ, its density function has the following form

$$f(Z \mid \mu, \sigma^2, \lambda) = 2N\!\left(z \mid \mu, \sigma^2\right) \cdot \Phi\!\left(\lambda \frac{z - \mu}{\sigma}\right)$$

where $N(z|\mu, \sigma^2)$ is the density of normal distribution with mean μ and variance σ^2, and Φ is the cumulative distribution of the standard normal distribution. One issue with this parameterization is that it does not control the mean of Z directly so that the zero-mean assumption of the random errors/effects in model (1)-(2) cannot be implemented easily. To solve this problem, we adopt an alternative parameterization in the simulations

$$Z \sim SN(\omega, \tau^2, \lambda)$$

$$f(Z \mid \omega, \tau^2, \lambda) = 2N\!\left(z \mid \mu, \sigma^2\right) \cdot \Phi\!\left(\lambda \frac{z - \mu}{\sigma}\right)$$

$$\mu = \omega - \sqrt{\frac{2}{\pi}} \cdot \frac{1}{\sqrt{\frac{1 + \lambda^2}{\lambda^2} - \frac{2}{\pi}}} \cdot \tau, \quad \sigma^2 = \frac{\tau^2}{1 - \frac{2}{\pi} \cdot \frac{\lambda^2}{1 + \lambda^2}}$$

where ω and τ^2 are the mean and variance of Z. Similarly, the skew-t distribution can be represented by

$$Z \sim ST(\omega, \tau^2, \lambda, v)$$

$$f(Z \mid \omega, \tau^2, \lambda, v) = \frac{2}{\sigma} t\!\left(z \mid \mu, \sigma^2, v\right) \cdot T\!\left(\lambda \frac{z - \mu}{\sigma} \sqrt{(v + 1)/\left(v + \left(\frac{z - \mu}{\sigma}\right)^2\right)} \,\middle|\, v + 1\right)$$

where v is the degree of freedom, and μ and σ^2 can be obtained using the same formulas as in the skew-normal distribution.

Using the skew-normal and the skew-*t* distribution, we can simulate different situations of non-normality by manipulating their parameters. Sampling from these distributions can be done using Markov chain Monte Carlo (MCMC) algorithms[34]. In our study, we use the slice sampler[35] through the *slicesample* function in Matlab to generate samples following the two distributions. Figure 4 shows the empirical distributions of examples of the simulated data, where $\omega=0$, $\tau^2=1$ and 100000 samples are generated in each case.
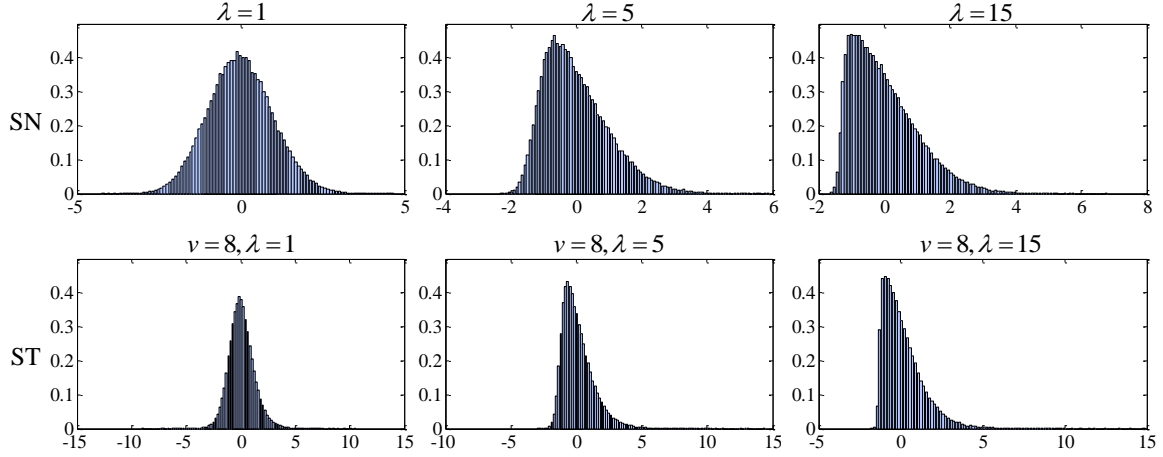
Figure 4. Normalized histograms of simulated data from skew-normal and skew-*t* distribution

4.2 Performance of data segmentation methods

In this study, data streams following skew-normal distribution ($\lambda=6$) and skew-*t* distribution ($\lambda=6$, $v=6$) are simulated. To obtain insight on the two data segmentation methods in multiple change point detection, a simple scenario is considered in which each data stream contains two equally-spaced change points (i.e., $K_1=100$, $K_2=200$, $m=300$) or in other words, three segments with equal length (100). The changes are either location or scale shifts. When location shifts occur, the scale parameters of the three segments take the same value ($\tau^2=1$), while their location parameters ω_1, ω_2, and ω_3 are different. Similarly, when scale shifts occur, the location parameters of the three segments are the same ($\omega=0$), while their scale parameters τ_1^2, τ_2^2 and τ_3^2 take different values. 4 cases are simulated under each type of shifts, which lead to a total of 8 cases. Table 1 summarizes the parameter settings and interpretations in these cases. Figure 5 shows an example of data streams generated in each case, where the solid line in each plot indicates the true value of the location parameter.

11

Table 1. Parameter settings in evaluating performance of data segmentation methods

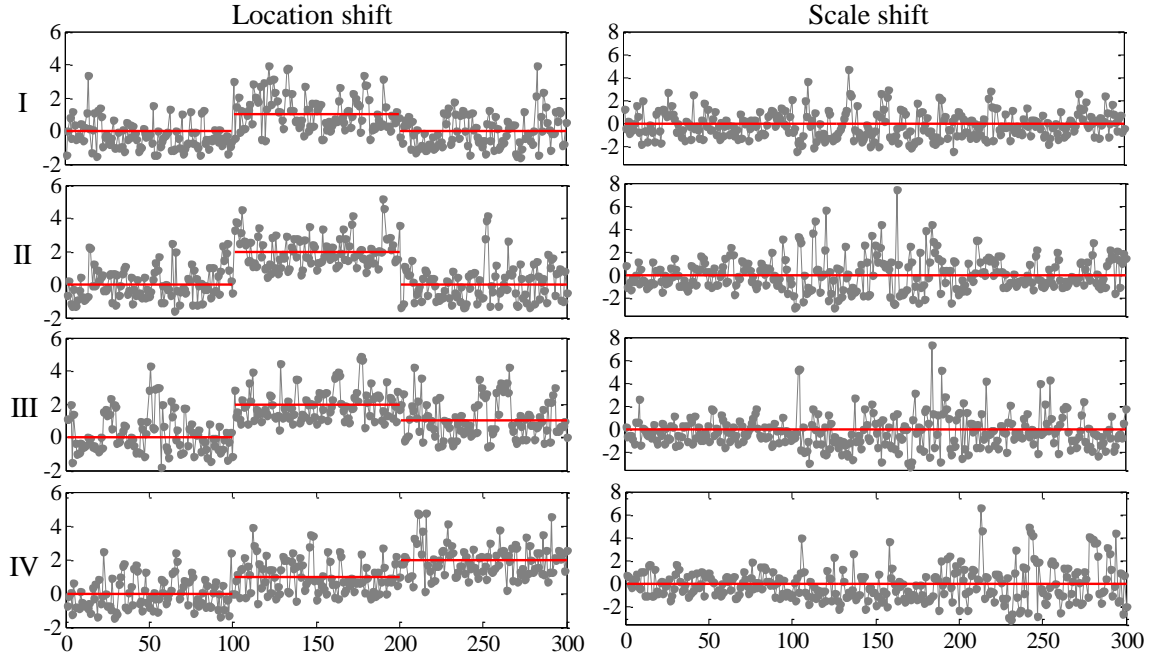| Case	Location shift			Scale shift			Interpretation
	ω_1	ω_2	ω_3	τ_1^2	τ_2^2	τ_3^2	
I	0	1	0	1	2.5	1	a small change, then back to in-control
II	0	2	0	1	4	1	a large change, then back to in-control
III	0	2	1	1	4	2.5	a large change, followed by a small change
IV	0	1	2	1	2.5	4	a small change, followed by a large change

Figure 5. Examples of data streams generated under each case listed in Table 1

Under each case listed in Table 1, 10000 data streams are simulated. The BS and the SS method are applied to each of the streams. In using the control charts in (4) and (5), a control limit with $ARL_{in\text{-}control} = 2000$ is applied. The performance of the two methods in each case is evaluated using the following measures:

R_{FA} = probability that more than 2 change points are detected

R_{MIS} = probability that only 1 change point or no change point is detected

R_{E1} = probability that the change point estimate is within 10% interval of K_1

R_{E2} = probability that the change point estimate is within 10% interval of K_2

Here the "10% interval" in the last two measures means the interval [90, 110] for K_1, and [190, 210] for K_2. The above performance measures essentially represent the false alarming rate, miss detection rate, and accuracy in estimating K_1 and K_2.

Table 2 gives the results on the performance measures in location shift detection. Figure 6 shows the corresponding distributions of change point estimates for skew-normal data. The change point estimates for skew-t data exhibit similar patterns. We find the following things from the results:

- **The performance of the BS and the SS method shows some common characteristics:** According to results in Table 2, both methods have lower miss detection rate and more accurate change point estimates when the location difference between the two sides of the change point is higher. From the upper panel of Figure 6, we can see that the estimates of K_1 in Case II and III have a sharper distribution than in other cases, meaning that the estimation is more accurate. This is because the difference in the locations at the two sides of K_1 is larger in these two cases. For the estimation of K_2, Case II performs the best as the location difference at the two sides of the change point in this case is larger than in other cases.

- **The two methods are different in two aspects:** (1) From Table 2, the BS method has much smaller false alarming rate and considerably larger miss detection rate than the SS method. This means that the BS method tends to miss some change points, while the SS method tends to detect some false change points. This is consistent to our intuitive understanding of these two methods given in Section 3.4: since the BS method works on a whole segment which contains more information, it is less likely to signal a false change point; but meanwhile it is more likely to miss some true change points as it can only pick one change point from the segment being inspected which may in fact contain multiple change points. In contrast, the SS method examines the data sequentially so that it is more likely to detect the true change points; but meanwhile it tends to generate more false alarms due to the limited information used especially at the beginning of each detection. (2) From Figure 6, we can see that the change point estimates from the two methods have similar distributions in general, with the mode of the SS method being slightly higher than the BS method. Overall we can say that they provide change point estimates of similar accuracy.

- **Comparing the skew-normal and skew-t data:** The results of the two distributions show similar patterns, but in most cases the skew-t data have higher false alarm rate and miss detection rate, and less accurate change point estimates than the skew-normal data.

13

Table 2. Performance of the BS and the SS method in detecting location shifts

|	Case	Binary segmentation				Sequential segmentation			
		R_{FA}	R_{MIS}	R_{E1}	R_{E2}	R_{FA}	R_{MIS}	R_{E1}	R_{E2}
SN	**I**	0.1049	0.0234	0.8081	0.8034	0.4722	0.0002	0.8711	0.8720
	II	0.1306	0	0.9841	0.9834	0.5075	0	0.9830	0.9969
	III	0.0724	0.0011	0.9797	0.8848	0.5068	0.0003	0.9849	0.8708
	IV	0.1655	0.0064	0.8281	0.8132	0.4481	0.0002	0.8680	0.8678
ST	**I**	0.1476	0.0938	0.6804	0.6820	0.6526	0.0025	0.8019	0.8075
	II	0.1885	0	0.9395	0.9403	0.6676	0	0.9477	0.9827
	III	0.1305	0.0106	0.9393	0.8026	0.6718	0.0005	0.9538	0.8030
	IV	0.1948	0.0452	0.7301	0.7219	0.6139	0.0040	0.8062	0.8020
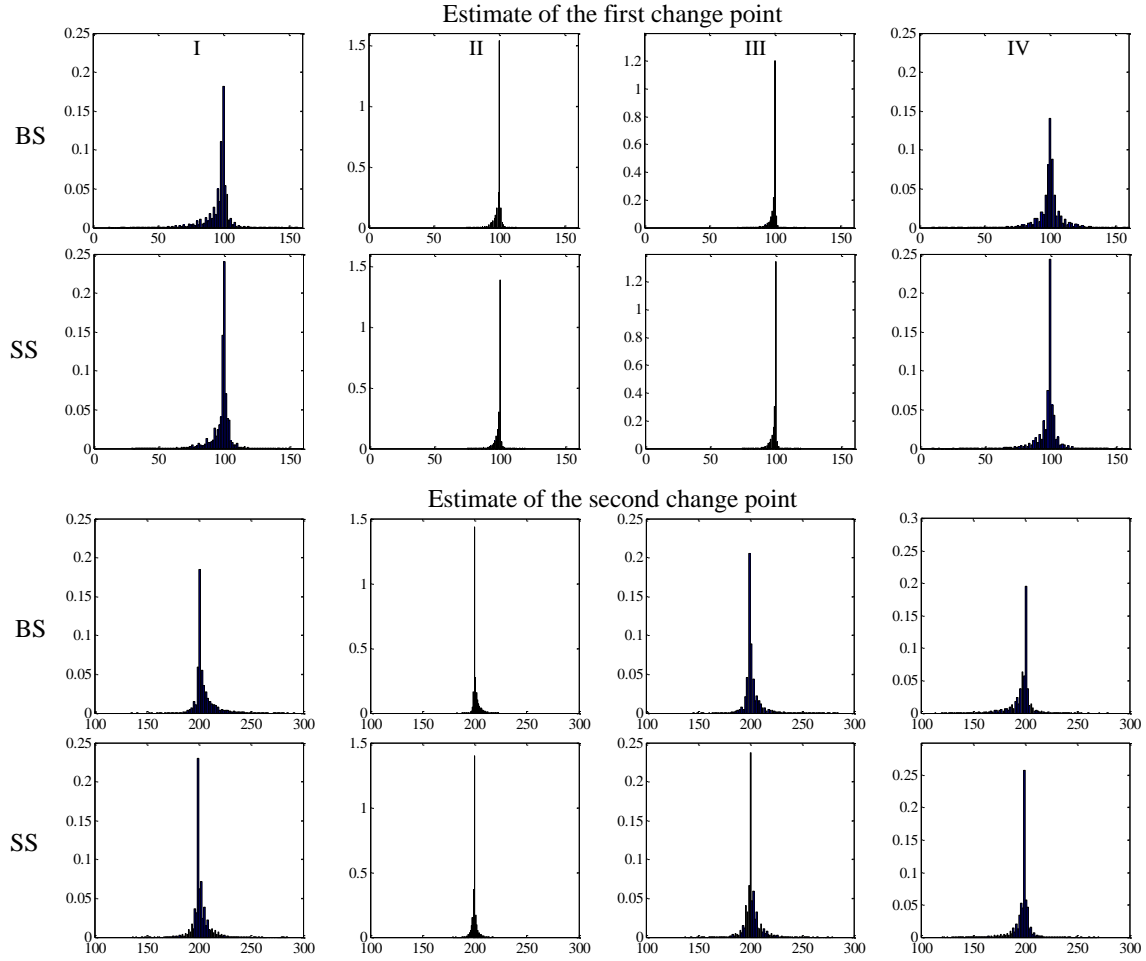
Figure 6. Normalized histograms of change point estimates under location shifts

The results on the performance measures in detecting scale shifts are given in Table 3, and the corresponding distributions of change point estimates for the skew-normal data are shown in Figure 7. In general, the performance of the two methods shows similar patterns as in the cases of location shifts. Both methods perform the best in Case II where the difference at the two sides

of the change points is larger than in other cases. The BS method has higher miss detection rate, while the SS method has higher false alarming rate. Both rates are larger than in the cases of location shifts. Correspondingly, the distribution of change point estimates has larger variance. This is because scale shifts are, in general, more difficult to detect than location shifts.

Table 3. Performance of the BS and the SS method in detecting scale shifts

|	Case	Binary segmentation				Sequential segmentation			
		R_{FA}	R_{MIS}	R_{E1}	R_{E2}	R_{FA}	R_{MIS}	R_{E1}	R_{E2}
SN	I	0.0702	0.5057	0.3470	0.3460	0.6097	0.0256	0.5994	0.5743
	II	0.1341	0.0757	0.7613	0.7618	0.6563	0.0007	0.8016	0.8137
	III	0.0411	0.5819	0.8383	0.1378	0.5213	0.2055	0.7737	0.2420
	IV	0.0280	0.7749	0.4958	0.1259	0.5918	0.1367	0.5895	0.2779
ST	I	0.0890	0.5385	0.3032	0.3072	0.6715	0.0324	0.5381	0.5319
	II	0.1670	0.1051	0.6943	0.6924	0.7195	0.0016	0.7672	0.7580
	III	0.0651	0.5775	0.7730	0.1312	0.6157	0.1547	0.7420	0.2575
	IV	0.0419	0.7679	0.4404	0.1296	0.6522	0.1223	0.5469	0.2855
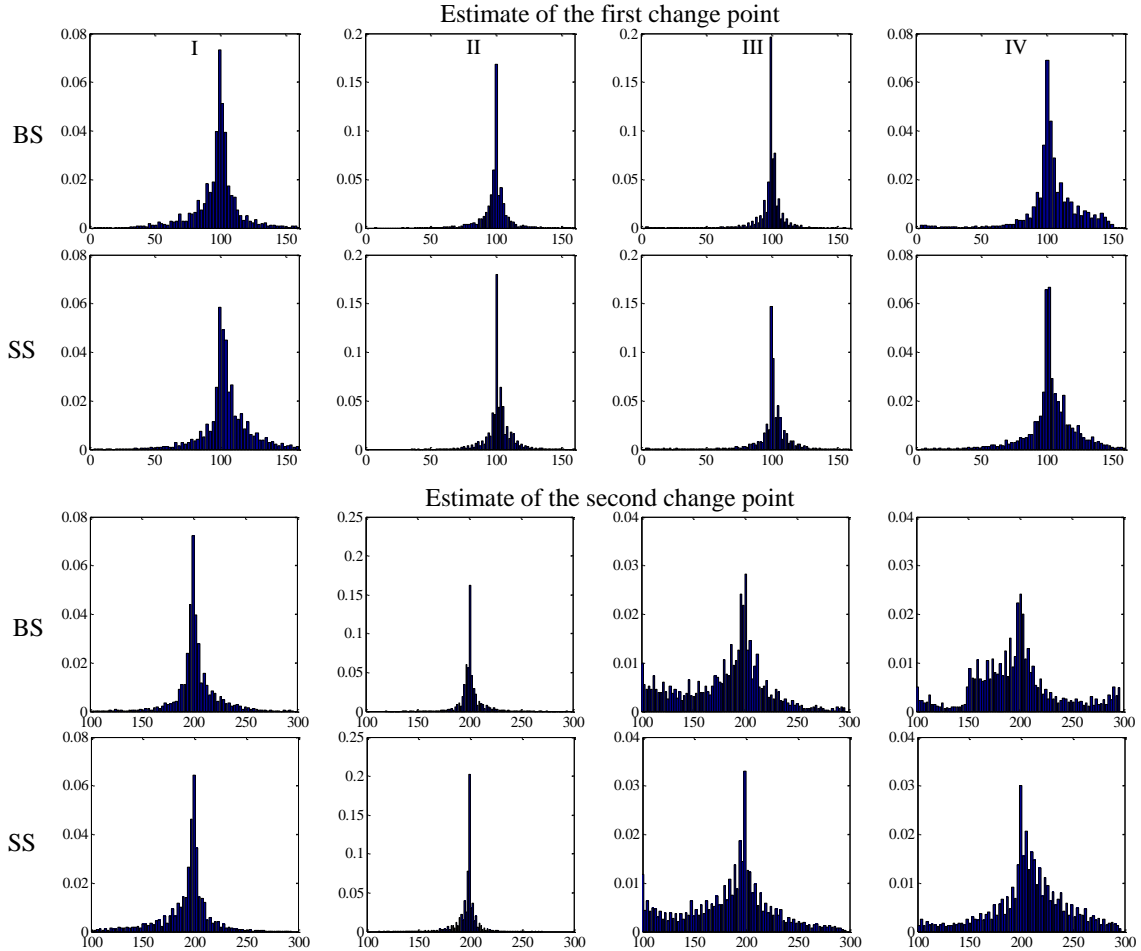
Figure 7. Normalized histograms of change point estimates under scale shifts

4.3 Properties of the proposed strategy for Phase I monitoring

In this study, we simulate streams of profile data following the regular polynomial model in (1) or the mixed-effect model in (2) with degree-2 polynomials and apply the proposed strategy for Phase I monitoring on each stream. Like the simulated data in Section 4.2, each stream contains three segments, and each segment contains 100 profiles following a different model. The random errors/random effects in the models are generated from skew-normal distributions. The in-control models are

Regular polynomial model:
$$y = \beta_2 x^2 + \beta_1 x + \beta_0 + \varepsilon$$
$$\beta_2 = \beta_1 = \beta_0 = 2, \quad \varepsilon \sim SN(\omega = 0, \tau^2 = 1, \lambda = 6)$$

Mixed-effect polynomial model:
$$y_i = \eta_2 x^2 + \eta_1 x + \eta_0 + \alpha_{2,i} x^2 + \alpha_{1,i} x + \alpha_{0,i} + \varepsilon$$
$$\eta_2 = \eta_1 = \eta_0 = 2$$
$$\alpha_{2,i} \sim SN(\omega = 0, \tau_2^2 = 1, \lambda = 6)$$
$$\alpha_{1,i} \sim SN(\omega = 0, \tau_1^2 = 1, \lambda = 6)$$
$$\alpha_{0,i} \sim SN(\omega = 0, \tau_0^2 = 1, \lambda = 6)$$
$$\varepsilon \sim SN(\omega = 0, \tau_\varepsilon^2 = 1, \lambda = 6)$$

where the explanatory variable x takes values [0, 0.1, 0.2,…., 3.0]. To be convenient, the change structure in Case I and II in Table 1 is applied to each data stream, that is, the first and third segments follow the above in-control model, while the second segment follows a different model. 6 cases are simulated considering different settings of the parameters of the second segment, which are listed in Table 4. Under each case, profile streams are generated and the proposed Phase I analysis is applied to each stream. The results of one typical example under each case are shown in Figure 8. In each plot of the figure, the left column displays the estimates of coefficients, while the right column displays the selected independent components. The estimated change points are marked in the plots of ICs.

Table 4. Parameter settings of the second segment in the simulated profile data

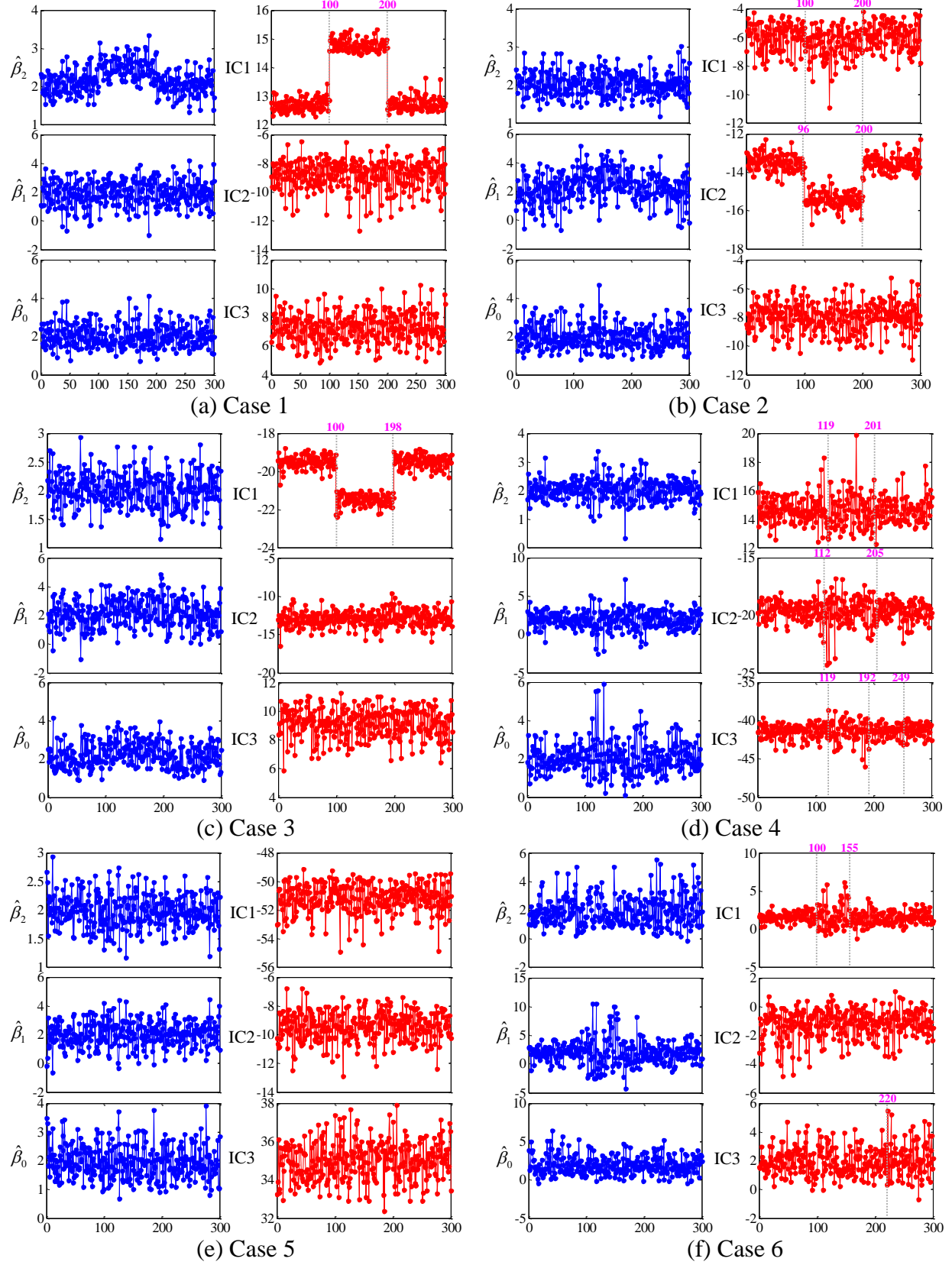| Case	Model	Parameters	Interpretation
1	Regular	$\beta_2 = 2.5$	Small location shift in quadratic coefficient
2	Regular	$\beta_1 = 3$	Mild location shift in linear coefficient
3	Regular	$\beta_1 = 2.5$, $\beta_0 = 2.5$	Small location shift in both linear coefficient and intercept
4	Regular	$\tau^2 = 2.5$	Small scale shift
5	Regular	$\lambda = 15$	Mild shift in skewness
6	Mixed-effect	$\tau_1^2 = 8$	Large shift in random-effect variance

Figure 8. Example of coefficient estimates and selected ICs under each case listed in Table 4

The results in Figure 8 can be summarized in the following aspects

- **The effect of ICA:** We can see that the shifts manifest themselves more clearly in the ICs than in the coefficient estimates. This is particularly the case in Figure 8(c) where the data contain two small location shifts. Little evidence of the shifts can be found in the coefficient estimates, while the evidence is quite apparent in the ICs. This validates the intrinsic capacity of ICA in manifesting the structure in the data. Another observation is that the shifts tend to appear in the first ICs, which implies the potential of ICA for data reduction when a large number of coefficients exist.

- **Change point estimation:** From Figure 8(a)-(c), it is seen that the change points of location shifts are estimated accurately. Not surprisingly, from Figure 8(d), we see that it is more difficult to estimate change points of scale shifts than location shifts. In Figure 8(f), due to the random effects of the coefficients in the mixed-effect model, estimation of the change points in scales becomes even more difficult. But according to our simulations not shown here, the accuracy in the estimation gets improved when the magnitude of the shift is larger. Finally, as shown in Figure 8(e), the two nonparametric control charts cannot detect shifts in skewness, which is reasonable as they are designed for location/scale shifts.

5. Case Study

In this study, the proposed Phase I analysis is applied to a set of optical profile data as shown in Figure 1. The data were from a large-scale low-E glass producer in the US. For confidentiality reasons, the name of the company and information of their products are not disclosed in this text. The data set consists of 314 optical profiles and each profile contains 30 data points corresponding to λ=[705nm, 710nm, …,850nm]. Before implementing the analysis, some preprocessing is done on the raw data. This includes the centering/scaling transformation of λ values, i.e., x=[λ–average(λ)]/150, which can improve the numerical properties of the fitting, and determining the appropriate degree of polynomials through fitting polynomial models to each profile and checking the residuals. As an example, Figure 9 shows the fitted models and resulting residuals for one profile. We can see that the residuals become very small and exhibit random patterns with equal variance when p=4. Therefore, we decide that the degree-4 polynomial model gives adequate fitting and will be used in the Phase I analysis.
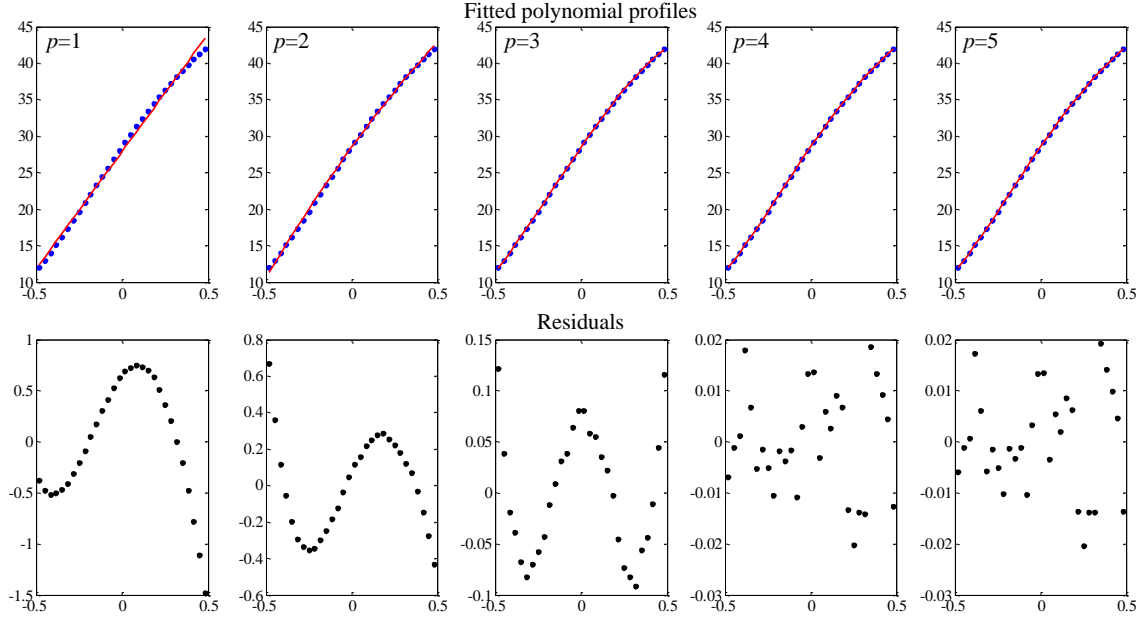
Figure 9. Example of fitted polynomial models and residuals

First, coefficient estimates are obtained from each profile, which are shown in Figure 10. The estimates consist of a considerable amount of extreme values, a sign of non-normality. It appears that multiple change points may exist in the data, and an apparent one of which occurs during profiles #200~#250. Then ICA is applied to these estimates. Figure 11 shows the resulting ICs. The apparent shift can be seen in the first IC, and there is also evidence of shifts in other ICs.

The BS and the SS method are applied to each IC. The estimates of change points are listed in Table 5. As expected, more change points are detected by the SS method, especially in detecting scale shifts. But the change point estimates from the two methods are very similar. For location shifts, multiple change points are detected including the apparent one (#232) in Figure 10. Fewer change points are detected for scale shifts. Particularly, only one change point is obtained by the BS method. Using the detected change points, the data are divided into multiple segments. Figure 12 and 13 show the segments based on the results of the SS method.

Table 5. Estimates of change points for each independent component

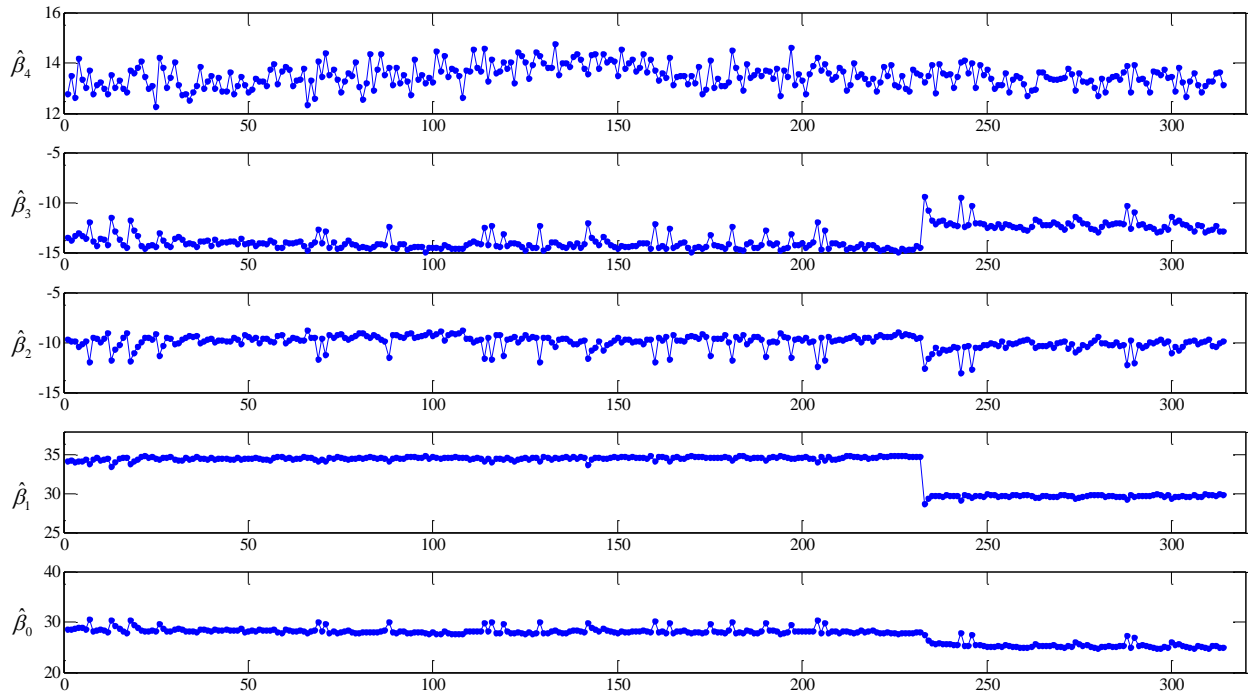| IC	Location shift detection		Scale shift detection	
	Binary seg.	Sequential seg.	Binary seg.	Sequential seg.
IC1	55, 100, 159, 232	55, 100, 159, 229	234	5,160, 232
IC2	89, 148, 232, 246	89, 148, 232, 246	N/A	36, 48,246
IC3	50, 140, 304	50, 140, 291	N/A	N/A
IC4	14, 100, 122, 162, 232, 248	14, 99, 122, 159, 232	N/A	118, 251
IC5	69, 128	13, 69, 128, 158, 247	N/A	159

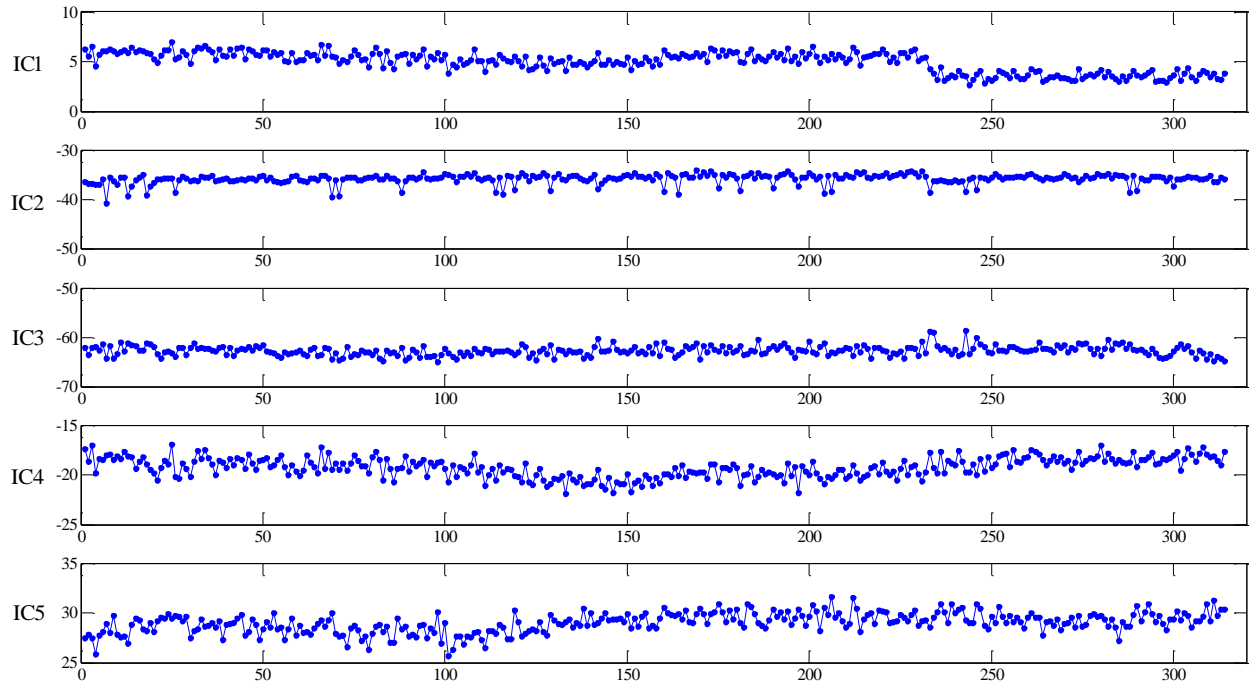Figure 10. Estimates of coefficients of degree-4 polynomial models

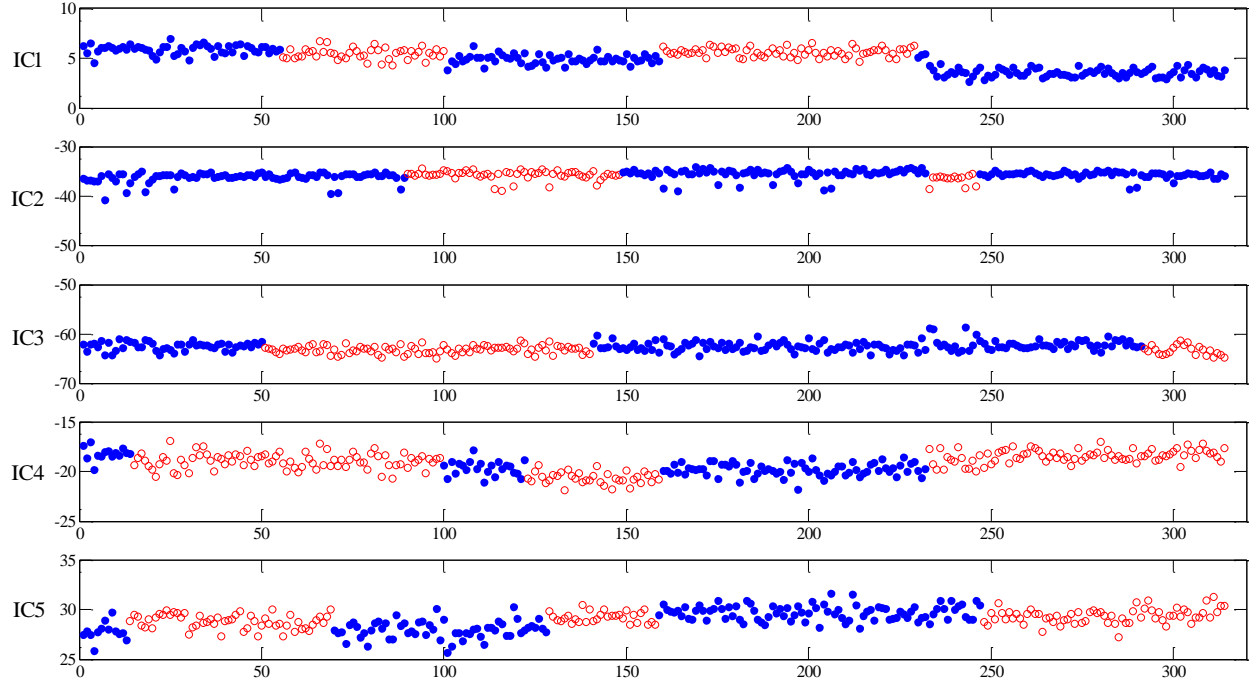Figure 11. Independent components obtained from the coefficient estimates

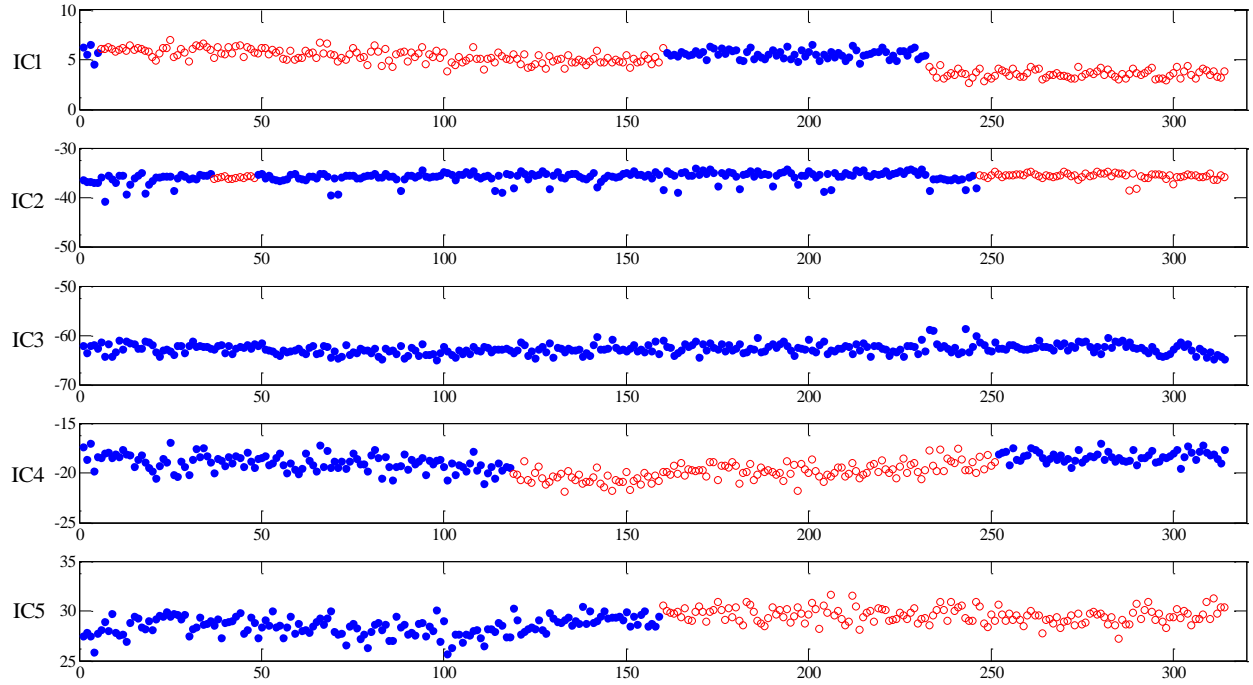Figure 12. Estimates of location change points using the sequential segmentation method

Figure 13. Estimates of scale change points using the sequential segmentation method

Based on the results of the two methods, two groups of profiles are identified which have constant location and scale. The two groups contain profiles #159~#232 and #246~#291, which are shown in Figure 14. They have apparently different shapes, indicating that the process

21

underwent a location shift. Degree-4 polynomial models are fitted for the two groups separately. Figure 15 shows the quantile-quantile (QQ) plots of the coefficient estimates of the first group. It is clear that the distribution of the estimates is not normal, which justifies the use of non-parametric change detection techniques.
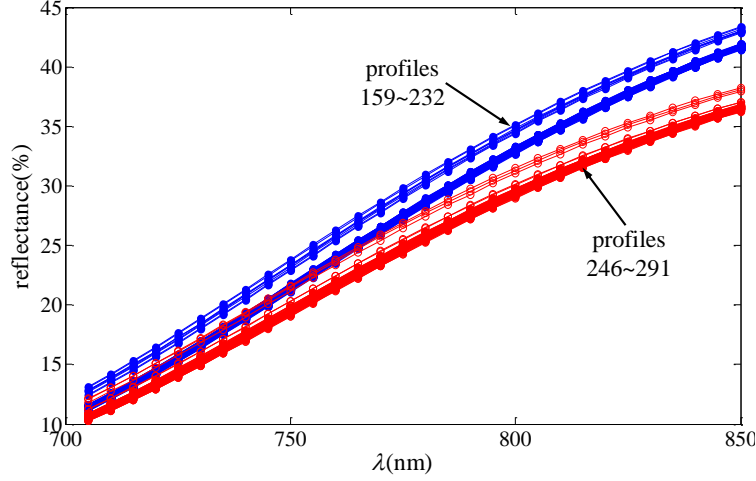

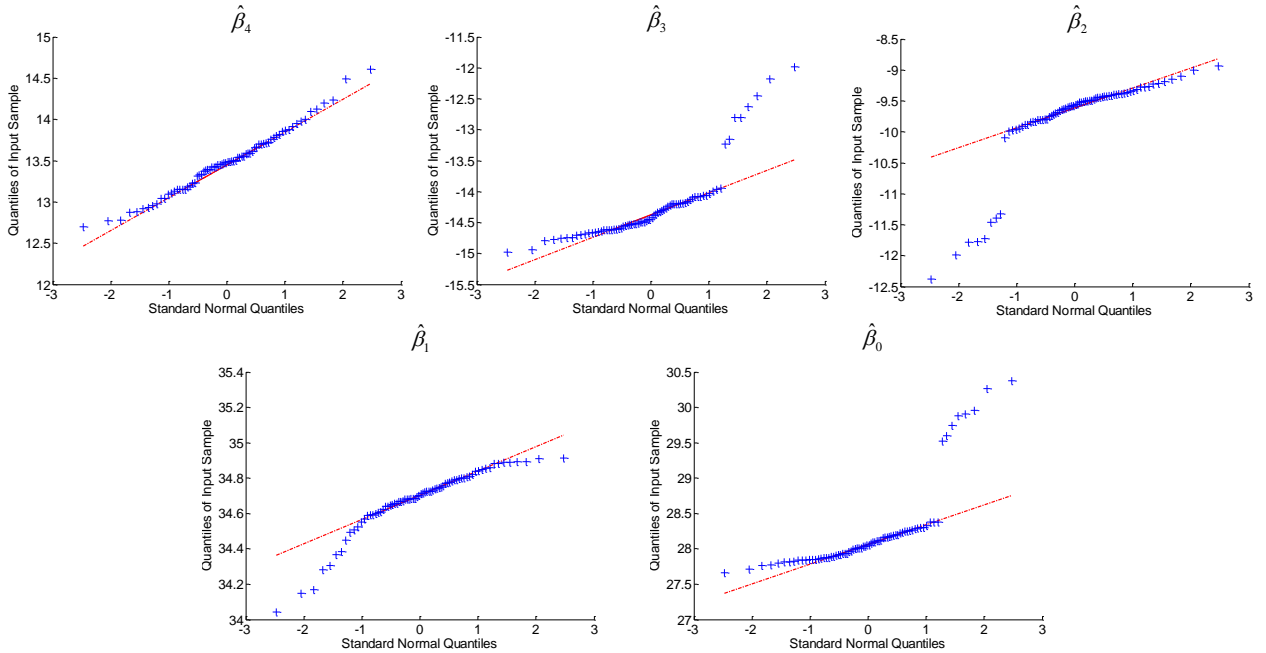Figure 14. The identified two groups of profiles

Figure 15. QQ-plots of the coefficient estimates of profiles #159~#232

We have consulted with the engineers on the findings in the Phase I analysis. After carefully reviewing the process history, they identified an abrupt change in the voltage/current of certain coating chambers which is likely to have caused the location shift between the two groups of

22

profiles in Figure 14. It is believed that such changes occur when the production switches to a different type of glass products. They also captured a number of small drifts in some process variables such as the oxygen/nitrogen flow in certain chambers. These drifts are likely to be the reason for other change points shown in Figure 12 and 13. Such drifts are in general difficult to control due to all sorts of random factors in the coating process.

6. Summary

This study proposes a strategy for Phase I monitoring of profile data under non-normality. The strategy contains three components: fitting appropriate models for profiles, independent component analysis on coefficient estimates, and change point detection on each independent component using nonparametric control charts. The performance of this strategy is studied through simulations on general classes of non-normal distributions. It is found that the use of ICA can reveal the structure in the data; between the two methods to detect multiple change points, binary segmentation has a lower false alarming rate, while sequential segmentation has a lower miss detection rate; the estimation of change points is more accurate when the difference in the location/scale at the two sides of the change point is larger. In the case study, the proposed strategy is applied to optical profiles from low-E glass manufacturing. A number of change points are detected, and two groups of profiles with constant location/scale are identified. Causes for the detected process shifts are also analyzed.

References

1. 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.
2. Kim, K., Mahmoud, M.A., and Woodall, W.H. (2003) On the monitoring of linear profiles. *Journal of Quality Technology*, **35**:317-328.
3. 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.
4. Kazemzadeh, R.B., Noorossana, R., and Amiri, A. (2009) Monitoring polynomial profiles in quality control applications. *International Journal of Advanced Manufacturing Technology*, **42**:703-712.
5. Kazemzadeh, R.B., Noorossana, R., and Amiri, A. (2010) Phase II monitoring of autocorrelated polynomial profiles in AR(1) processes. *Transaction E: Industrial Engineering*, **17**(1): 12-24.

6. Amiri, A., Jensen, W.A., and Kazemzadeh, R.B. (2009) A case study on monitoring polynomial profiles in the automotive industry. *Quality and Reliability Engineering International*, **26**:509-520.

7. Walker, E., and Wright, S.P. (2002) Comparing curves using additive models, *Journal of Quality Technology*, **34**:118-129.

8. Jensen, W. A., Birch, J. B., and Woodall, W. H. (2008) Monitoring Correlation within Linear Profiles Using Mixed Models. *Journal of Quality Technology*, **40**(2):167-183.

9. Jensen, W.A. and Birch, J.B. (2009) *Profile monitoring via nonlinear mixed models. Journal of Quality Technology*, **41**(1):18–34.

10. 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**:925-941.

11. Noorossana, R., Saghaei, A., and Amiri, A. (2011) *Statistical analysis of profile monitoring*. John Wiley & Sons, Hoboken, NJ.

12. Woodall, W.H. (2007) Current research on profile monitoring. *Produção*, **17**(3), 420-425.

13. Mahmoud, M.A., and Woodall, W.H. (2004) Phase I analysis of linear profiles with calibration applications. *Technometrics*, 46(4):380-391.

14. Mahmoud, M.A., Parker, P.A., Woodall, W.H., and Hawkins, D.M. (2007) A change point method for linear profile data. *Quality and Reliability Engineering International*, 23:247-268.

15. Kazemzadeh, R.B., Noorossana, R., and Amiri, A. (2008) Phase I monitoring of polynomial profiles. *Communications in Statistics-Theory and Methods*, **37**:1671-1686.

16. Arasteh, D., Carmody, J., Lee, E.S., and Selkowitz, S. (2004) *Window systems for high-performance buildings*, W. W. Norton & Company, New York, NY.

17. Carmody, J., Selkowitz, S., and Heschhong, L. (1996) *Residential windows: a guide to new technologies and energy performance*. W. W. Norton & Company, New York, NY.

18. Frost, K., Arasteh, D., and Eto, J. (1993) Savings for energy efficient windows: current and future savings from new fenestration technologies in the residential market. *Lawrence Berkeley Laboratory Report 33965*, Lawrence Berkeley Laboratory.

19. Williams, J.D., Birch, J.B., Woodall, W.H., and Ferry, N.M. (2007) Statistical monitoring of heteroscedastic dose-response profiles from high-throughput screening. *Journal of Agricultural, Biological and Environmental Statistics*, 12:216-235.

20. Vaghefi, A., Tajbakhsh, S.D., and Noorossana, R. (2009) Phase II monitoring of nonlinear profiles. *Communications in Statistics: Theory and Methods*, **18**:1834-1851.

21. Noorossana, R., Vaghefi, A., and Dorri, M. (2011) Effect of non-normality on the monitoring of simple linear profiles. *Quality and Reliability Engineering International*, **27**:425-436.

22. Zou, C., and Tsung F. (2011) A multivariate sign EWMA control chart. *Technometrics*, **53**(1):84-97.

23. Qiu, P., and Hawkins, D. (2001) A Rank-based multivariate CUSUM procedure. *Technometrics*, **43**(2):120-132.

24. Qiu, P., and Hawkins, D. (2003) A nonparametric multivariate cumulative sum procedure for detecting shifts in all directions. *Journal of the Royal Statistical Society. Series D*, **52**(2):151-164.

25. Hyvarinen, A., Harhunen, J., and Oja, E. (2001) *Independent component analysis*. Wiley, New York, NY.

26. Ding, Y., Zeng, L., and Zhou, S. (2006) Phase I analysis for monitoring nonlinear profiles in manufacturing processes. *Journal of Quality Technology*, **38**(3):199-216.

27. Jones, L., and Woodall, W. (1998) The performance of Bootstrap control charts. *Journal of Quality Technology*, **30**:362-375.

28. Gordon, L., and Pollak, M. (1994) An efficient sequential nonparametric scheme for detecting a change of distribution. *The Annals of Statistics*, **22**:763-804.

29. Hackl, P., and Ledolter, J. (1991) A control chart based on ranks. *Journal of Quality Technology*, **23**:117-124.

30. Hawkins, D.M., and Deng, Q. (2010) A nonparametric change-point control chart. *Journal of Quality Technology*, **42**(2):165-173.

31. Ross, G.J., Tasoulis, D.K., and Adams, N.M. (2011) Nonparametric monitoring of data streams for changes in location and scale. *Technometrics*, **53**(4):379-389.

32. Azzalini, A. (1985) A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, **12**:171-178.

33. Azzalini, A., and Genton, M.G. (2008) Robust likelihood methods based on the skew-*t* related distributions. *International Statistical Review*, **76**(1):106-129.

34. Robert, C.P., and Casella, G (2004) *Monte Carlo statistical methods*, *2nd edition*, Springer, New York, NY.

35. Neal, R. M. (2003) "Slice Sampling", *The Annals of Statistics*, **31**:705-741.