Investigating Effect of Autocorrelation on Monitoring Multivariate Linear Profiles

P. Soleimani & R. Noorossana *

Paria Soleimani, Department of Industrial Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran
Rasool Noorossana, Industrial Engineering Department, Iran University of Science and Technology, Tehran, Iran

KEYWORDS
Multivariate linear profiles, Autocorrelation, Time series modeling, Average Run Length

ABSTRACT
Profile monitoring in statistical quality control has attracted attention of many researchers recently. A profile is a function between response variables and one or more independent variables. There have been only a limited number of researches on monitoring multivariate linear profiles. Indeed, monitoring correlated multivariate profiles is a new subject in the field of statistical process control. In this paper, we investigate the effect of autocorrelations in monitoring multivariate linear profiles in phase II. The effect of three main models namely AR(1), MA(1), and ARMA(1,1) on the methods of multivariate linear profile monitoring is evaluated and compared by using simulation study and average run length criteria. Results indicate that autocorrelation affects performance of the existing methods significantly.


1. Introduction
Profile monitoring is a relatively new quality control concern with many applications. Many authors have recently investigated issues related to profile monitoring. Kang and Albin [1] and Kim et al. [2] introduced methods to monitor simple linear profiles. Zou et al. [3] and Mahmoud et al. [4] considered change point methods in profile monitoring. Kazemzade et al. [5] studied polynomial profiles. Zou et al. [6] combined multivariate exponentially weighted moving average procedure with a generalized likelihood ratio test based on nonparametric regression to monitor nonlinear profiles. Also nonlinear profiles monitoring was discussed by researchers including Ding et al. [7], Moguerza et al. [8], Williams et al. [9], and Vaghefi et al.[10]. Noorossana et al. [11, 12], Zou et al. [13], and Eyyazian et al. [14] proposed methods to monitor multivariate linear profiles. Noorossana et al. [15] showed the effect of non–normality on the monitoring of simple linear profiles. Several authors including Jensen et al. [16], Noorossana et al. [17, 18, and 19], Jensen and Birch [20], Soleimani et al. [21, 22, 23, and 24], Kazemzadeh et al. [25] addressed issues related to autocorrelation in linear, non-linear, and polynomial profiles. Soleimani and Noorossana [26, 27] proposed methods to consider within and between profile autocorrelation in multivariate linear profiles in phase II. Recently, new topics such as wavelet filtering, high dimensional control chart, and roundness profile were studied by Chang et al. [28], Chen et al. [29], and Pacella et al. [30], respectively. Independence of within or between error terms is one of the basic assumptions in most of the profile monitoring methods. However, in certain situation this assumption can be violated easily.

* Corresponding author: Rasool Noorossana
Email: rassoul@iust.ac.ir
In this paper, we investigate the effect of autocorrelation within multivariate simple linear profiles in phase II. We consider the multivariate simple linear profile model presented by Noorossana et al. [12] or

\[ Y_k = X\beta + E_k \]  

where \( Y_k \) is an \( n \times 1 \) matrix of response variables for the \( k \)th sample, \( X \) is an \( n \times 2 \) matrix of independent variable, \( \beta \) is a \( 2 \times 1 \) matrix of known regression parameters, and \( E_k \) is an \( n \times 1 \) matrix of error terms which follows a multivariate normal distribution with mean vector zero and known covariance matrix \( \Sigma \). In this study, we consider the well known least squares estimator of \( \beta \) defined as:

\[ \hat{\beta} = (X^T X)^{-1} X^T Y_k \quad k = 1, 2, ... \]  

Section 2 presents a review on multivariate simple linear profile monitoring methods in phase II. Autocorrelated models are presented in Section 3. In Section 4, effects of autocorrelation on the average run length performance of the proposed models are investigated. Section 5 summarizes our concluding remarks.

2. The Multivariate Simple Linear Profile Monitoring Methods

The three methods proposed by Noorossana et al. [12] for monitoring multivariate simple linear profiles in phase II are as follows.

The first method is based on MEWMA control chart. The coefficient vector for \( \hat{\beta}_k \) can be written as:

\[ \hat{\beta}_k^T = (\hat{\beta}_{1k}, \hat{\beta}_{2k}, ..., \hat{\beta}_{1k}, \hat{\beta}_{2k}) \]  

For an in control process, \( \hat{\beta}_k^T \) is a multivariate normal vector with known mean vector defined as:

\[ \mu = (\mu_{1k}, \mu_{2k}, ..., \mu_{1k}, \mu_{2k}) \]  

and a 2(1) covariance matrix \( \Sigma_k \) with the following correlation structure between its elements:

\[ \text{Cov}(\hat{\beta}_{uk}, \hat{\beta}_{vk}) = \sigma_{uv} \left( \frac{1}{n} + \frac{1}{s_{uk}} \right) \]  

where \( \sigma_{uv} \) and \( s_{uk} \) are the \( u^{th} \) row and the \( v^{th} \) column of the covariance matrix \( \Sigma \) and correlation matrix \( R \), respectively, where \( r_{uv} = \sigma_{uv}/\sigma_{u}\sigma_v \).

The multivariate exponentially weighted moving average (MEWMA) vector is defined as:

\[ z_{k,\beta} = \omega(\hat{\beta}_k - \beta) + (1 - \omega)z_{k-1,\beta} \]  

where \( z_{k,\beta} \) is a multivariate normal random vector with zero mean vector and known covariance matrix \( \Sigma_{z,\beta} = (\omega/2 - \omega)\Sigma \). For monitoring the coefficients vector, the chart statistic is defined as (Lowry et al. [31]).

\[ T_{z(k,\beta)}^2 = z_{k,\beta} \Sigma_{z,\beta}^{-1} z_{k,\beta}^T \]  

when \( T_{z(k,\beta)}^2 > h_\beta \), this chart gives an out of control signal where \( h_\beta (\beta > 0) \) is chosen to have a specific in-control average run length (ARL).

The second method referred to as MEWMA \( \xi^2 \) uses the MEWMA vector for monitoring mean vector of error terms, \( \xi_k = (\xi_{1k}, \xi_{2k}, ..., \hat{\xi}_{1k}, \hat{\xi}_{2k}) \) where \( \Sigma_{z,\beta} = (\omega/2 - \omega)\Sigma \). For monitoring the vector of error, the chart statistic is defined as:

\[ T_{z(k,\beta)}^2 = z_{k,\beta} \Sigma_{z,\beta}^{-1} z_{k,\beta}^T \]  

where \( T_{z(k,\beta)}^2 > h_\beta \), this chart gives an out of control signal where \( h_\beta (\beta > 0) \) is selected to have a specific in-control ARL. A chi-square chart with statistic \( \chi^2 = \sum_{i=1}^{n} \xi_i^2 \) where \( \xi_i = e_i + \Sigma_{i}^{1/2} \Sigma_i^{1/2} e_i \) is used to monitor variation. The upper control limit is \( \text{UCL} = \chi_{\alpha}^{2n} \).

In the third method, in order to make intercepts vector independent of the slopes vector, they coded the \( \hat{\beta}_k \) th observation in the \( k \)th sample as:

\[ \gamma_{ik} = \beta_{ik}' + \epsilon_{ik} \]  

where \( \epsilon_{ik} = (x_1 - \bar{x}_1, \beta_{ik}' \beta_{ik}') \). When process is in control, \( \tilde{\beta}_{ik} \) and \( \hat{\beta}_{ik} \) are multivariate normal random vectors with mean vectors \( \beta_{ik}' \) and covariance matrices \( \Sigma_{ik} = n^{-1} \Sigma \) and \( \Sigma_{ik}^{-1} = (s_{ik})^{-1} \Sigma \), respectively. For monitoring the intercept vector, the chart statistic is given as:
\[ T_k^2 = z_k \sum_{i=1}^{n-1} z_i^T \]  
(9)

where 
\[ z_k = \omega(\beta_1 - \beta_1') + (1 - \omega)z_{k-1} \]  
and 
\[ \sum_{ik} = (\omega [2 - \omega]) \sum_{i=1}^{n-1} e_{ik} \]  
for monitoring the slope vector, the chart statistic is defined as:

\[ T_k^2 = z_{sk} \sum_{i=1}^{n-1} z_{sk}^T \]  
(10)

where 
\[ z_{sk} = \omega(\beta_1 - \beta_1') + (1 - \omega)z_{sk-1} \]  
and 
\[ \sum_{sk} = (\omega [2 - \omega]) \sum_{i=1}^{n-1} e_{sk} \]  
for monitoring profile variability and MEWMA statistic defined as (Crowder and Hamilton [32]).

\[ z_{sk} = \max \{ \omega \ln(\hat{\theta}_k^2) + (1 - \omega)z_{sk-1}, np \} \]  
(11)

The MEWMA-3 control chart gives an out of control signal when 
\[ T_k^2 > h_k \]  
or 
\[ T_k^2 > h_k \]  
or 
\[ z_{sk} > h_k \]  
where \( h_k \) and \( h_{sk} \) are chosen to achieve a specified in-control ARL.

### 3. Auto Correlated Multivariate Simple Linear Profile models

In order to show the effect of autocorrelation on the performance of multivariate profile monitoring, we consider three well known time series models, namely first order autoregressive model or AR(1), first order moving average model or MA(1), and first order autoregressive-moving average, ARMA(1,1). We consider a multivariate simple linear profile when an AR(1) autocorrelation structure exists in the error terms. Hence, for the \( k^{th} \) sample we have:

\[ y_{ik} = \beta_0 + x_i \beta_1 + \epsilon_{ik} \]  
\[ i = 1, 2, ..., n, \quad k = 1, 2, ... \]

Where:

\[ \epsilon_{ik} = \epsilon_{i-1k} + u_{ik} \]  
(12)

In addition, a multivariate simple linear profile model when the error terms have a MA(1) autocorrelation structure is:

\[ y_{ik} = \beta_0 + x_i \beta_1 + \epsilon_{ik} \]  
\[ i = 1, 2, ..., n, \quad k = 1, 2, ... \]

Where:

\[ \epsilon_{ik} = u_{ik} - u_{i-1k} \]  
(13)

Also, we investigate a multivariate simple linear profile model with ARMA(1,1) structure as follows:

\[ y_{ik} = \beta_0 + x_i \beta_1 + \epsilon_{ik} \]  
\[ i = 1, 2, ..., n, \quad k = 1, 2, ... \]

where 
\[ \epsilon_{ik} = \epsilon_{i-1k} + u_{ik} - u_{i-1k} \]  
(14)

In the above equation, \( \phi \) and \( \theta \) define the coefficient matrices. For the sake of simplicity, we consider them as diagonal matrices (\( k \times l \)) and diagonal elements \( (\phi, \theta) \) are the same for each matrix. The vector \( u_{ik} \) consists of normal random variables with zero mean and covariance matrix \( \Sigma \).

### 4. The Effect of Autocorrelation on ARL Performance

In this section, we investigate effect of autocorrelation on the ARL performances of the three methods discussed in Section 2 and the three models presented in Section 3. We consider the profiles used by Noorossana et al. [12] defined as:

\[ Y_1 = 3 + 2x + \epsilon_y \]  
\[ Y_2 = 2 + 1x + \epsilon_y \]  
(15)

where \( x = [2 4 6 8] \) is independent variables vector, \( \sigma_1^2 = \sigma_2 = 1 \) and \( \rho_1 = 0.5 \). In our study, we consider the effect of weak correlation \( (\phi = 0.1) \) and strong correlation \( (\phi = 0.9) \). It is clear when \( \phi \) is equal to \( \theta \), the autocorrelation structure leads to the independent situation. The results are based on 5,000 simulation runs. We used the original limits for the three methods leading to an overall in-control ARL of 200.

We evaluate the different shifts in intercept, slope and standard deviation of the profile (Eq.15) for MEWMA method. Table 1 shows the ARL performance when \( \beta_{01} \) shifts to \( \beta_{01} + \beta_{01} \sigma_1 \). Table 2 and 3 summarize the results for shift in \( \beta_{11} \) and \( \sigma_1 \) respectively.

<table>
<thead>
<tr>
<th>Model</th>
<th>( \lambda_0 )</th>
<th>0</th>
<th>0.2</th>
<th>0.4</th>
<th>0.6</th>
<th>0.8</th>
<th>1</th>
<th>1.2</th>
<th>1.4</th>
<th>1.6</th>
<th>1.8</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independent</td>
<td>( \phi = 0, \theta = 0 )</td>
<td>200.0</td>
<td>53.9</td>
<td>14.4</td>
<td>7.3</td>
<td>4.9</td>
<td>3.7</td>
<td>3.0</td>
<td>2.5</td>
<td>2.2</td>
<td>2.0</td>
<td>1.9</td>
</tr>
<tr>
<td>AR(1)</td>
<td>( \phi = 0.9 )</td>
<td>5.7</td>
<td>5.6</td>
<td>4.9</td>
<td>4.2</td>
<td>3.6</td>
<td>3.3</td>
<td>2.7</td>
<td>2.4</td>
<td>2.1</td>
<td>1.9</td>
<td>1.8</td>
</tr>
<tr>
<td>MA(1)</td>
<td>( \theta = 0.1 )</td>
<td>117.4</td>
<td>40.2</td>
<td>13.3</td>
<td>7.1</td>
<td>4.8</td>
<td>3.7</td>
<td>3.0</td>
<td>2.6</td>
<td>2.3</td>
<td>2.1</td>
<td>1.9</td>
</tr>
<tr>
<td>ARMA(1,1)</td>
<td>( \phi = 0.9, \theta = 0.1 )</td>
<td>135.4</td>
<td>53.4</td>
<td>15.0</td>
<td>7.2</td>
<td>4.8</td>
<td>3.6</td>
<td>2.9</td>
<td>2.5</td>
<td>2.2</td>
<td>2.0</td>
<td>1.9</td>
</tr>
<tr>
<td></td>
<td>( \phi = 0.1, \theta = 0.9 )</td>
<td>258.4</td>
<td>61.9</td>
<td>15.2</td>
<td>7.4</td>
<td>5.0</td>
<td>3.7</td>
<td>3.0</td>
<td>2.6</td>
<td>2.3</td>
<td>2.1</td>
<td>1.9</td>
</tr>
</tbody>
</table>

| Tab. 1. The average run length results for MEWMA method when \( \beta_{01} \) shifts to \( \beta_{01} + \lambda_0 \sigma_1 \) |
The results for the three monitoring methods, for the case of a shift in intercept, are summarized in Table 4. The order of the ARL is significantly affected when autocorrelation is considered. Figure 1 shows the results in Table 1 graphically. These results indicate that the in-control ARL is significantly affected when autocorrelation is present. However, as the shift size increases, the out-of-control ARL approaches the out-of-control ARL of no autocorrelation case. The results for the three methods are summarized in Table 4. The order of models name shows the severity of the effect of correlation.

The following results could be concluded from Table 4:

1. In general, positive autocorrelation reduces the in-control ARL or equivalently increases the false alarm rate.
2. According to the simulation results, among the considered correlation structures, AR(1) and ARMA(1,1) have more considerable effects on the performance of monitoring methods.
3. In general, by increasing the value of shift size, performance of the three correlation models become similar and correlation effects turn to be negligible.
4. In all the three monitoring methods, for the case of MA(1) model with weak correlation and small shifts, we can see an increase in ARL.

![Fig. 1. The ARL comparison for shift in intercept in strong and weak correlation conditions.](image-url)
Tab. 4. Comparison of the three correlation models for shifts in the intercept, slope, and standard deviation.

<table>
<thead>
<tr>
<th>Method</th>
<th>Correlation</th>
<th>Intercept</th>
<th>Slope</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEWMA</td>
<td>Strong</td>
<td>AR(1), ARMA(1,1), MA(1)</td>
<td>AR(1), ARMA(1,1), MA(1)</td>
<td>AR(1), ARMA(1,1), MA(1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>For large shifts ($\geq 1$), the</td>
<td>For large shifts ($\geq 0.15$), the</td>
<td>For large shifts ($\geq 1.6$), correlation effect for MA is negligible and the performances of AR and ARMA models are similar.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>performances of three models are</td>
<td>performances of three models are</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>similar and correlation effect is</td>
<td>similar and correlation effect is</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>negligible.</td>
<td>negligible.</td>
<td></td>
</tr>
<tr>
<td>MEWMA$\chi^2$</td>
<td>Strong</td>
<td>AR(1), ARMA(1,1), MA(1)</td>
<td>AR(1), ARMA(1,1), MA(1)</td>
<td>AR(1), ARMA(1,1), MA(1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>for small shifts ($&lt; 1$), MA increasing ARL.</td>
<td>for large shifts ($\geq 0.15$), MA</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>AR(1), ARMA(1,1), MA(1)</td>
<td>increasing ARL.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>For large shifts ($\geq 1$), the</td>
<td>For large shifts ($\geq 0.15$), the</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>performances of three models are</td>
<td>performances of three models are</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>similar and correlation effect is</td>
<td>similar and correlation effect is</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>negligible.</td>
<td>negligible.</td>
<td></td>
</tr>
<tr>
<td>MEWMA-3</td>
<td>Strong</td>
<td>AR(1), ARMA(1,1)</td>
<td>AR(1), ARMA(1,1)</td>
<td>AR(1), MA(1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>for small shifts ($&lt; 0.8$), MA increasing ARL.</td>
<td>for large shifts ($\geq 0.15$), MA</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>AR(1), ARMA(1,1)</td>
<td>increasing ARL.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>For moderate shifts ($\geq 0.6$), the</td>
<td>For large shifts ($\geq 0.15$), the</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>performances of three models are</td>
<td>performances of three models are</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>similar and for large shifts ($\geq 1.6$)</td>
<td>similar and correlation effect is</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>correlation effect is negligible.</td>
<td>negligible.</td>
<td></td>
</tr>
</tbody>
</table>

5. Conclusion

In this paper, the effect of three well known time series models namely AR(1), MA(1), and ARMA(1,1) were investigated on the performance of three multivariate linear profile monitoring methods. We considered three common methods referred to as MEWMA, MEWMA-3, and MEWMA$\chi^2$ for monitoring multivariate linear profiles in phase II. Simulation results indicate that autocorrelation affects ARL performance of the three monitoring methods significantly.

References


International Journal of Industrial Engineering & Production Research, September 2012, Vol. 23, No. 3
Average Control Chart”, Technometrics Vol. 34, 1992, pp. 46-53
