

# Comparison of Phase II Control Charts Based on Variable Selection Methods

Giovanna Capizzi and Guido Masarotto

**Abstract** In recent years, control charts based on variable selection (VS) algorithms have been suggested for monitoring multivariate data. These charts share the common idea that process faults usually affect a small fraction of the monitored quality characteristics. Thus, VS methods can be used to identify the subset of the variables for which the shift may have occurred. However, the suggested VS-based control charts differ in many aspects such as the particular VS algorithm and the type of control statistic. In this paper, we compare VS-based control charts in various out-of-control scenarios characterizing modern manufacturing environments such as high-dimensional data, profile, and multistage process monitoring. The main aim of this paper is to provide practical guidelines for choosing a suitable VS-based monitoring scheme.

**Keywords** Generalized least squares • LAR • LASSO • Multivariate EWMA

## 1 Introduction

Nowadays, the performance of modern processes depends on several related quality characteristics. The statistical monitoring of “high-dimensional” processes is known as multivariate statistical process control (MSPC, see Bersimis et al. 2007, for a comprehensive review of the MSPC literature). A critical task for an MSPC control scheme is assessing whether the multidimensional process is in-control (IC) or not. Although it is unlikely that all the quality characteristics shift simultaneously, it is more common that only a subset of variables experiences abnormal changes. Thus, it could be more efficient to monitor only the potential out-of-control (OC) variables, which, however, are not known in advance. Thus, recent developments in the MSPC framework propose using variable selection (VS) algorithms to identify the suspected variables and then charting only these characteristics to test whether the multidimensional process is in-control or not

---

G. Capizzi (✉) • G. Masarotto  
Department of Statistical Sciences, University of Padua, Padua, Italy  
e-mail: [capizzi@stat.unipd.it](mailto:capizzi@stat.unipd.it); [masarotto@stat.unipd.it](mailto:masarotto@stat.unipd.it)

(see Wang and Jiang 2009; Zou and Qiu 2009; Zou et al. 2010; Capizzi and Masarotto 2011; Jiang et al. 2012). These VS-based approaches seem attractive since they offer a very satisfactory performance in OC scenarios involving a shift in one, two, . . . , all the monitored quality characteristics. They are also coupled with diagnostic tools to accurately identify the variables responsible for the change.

These recent proposals combine different multivariate control charts with different VS procedures. A “forward” selection algorithm (FVS) has been combined with a Shewhart-type and a multivariate EWMA (MEWMA) by Wang and Jiang (2009) and Jiang et al. (2012), respectively. Other VS algorithms such as Least Absolute Shrinkage and Selection Operator (LASSO, see Tibshirani 1996) and Least Angle Regression (LAR, see Efron et al. 2004) have been proposed combined with an MEWMA-based control chart, by Zou and Qiu (2009), Zou et al. (2010) and Capizzi and Masarotto (2011), respectively. The suggested monitoring schemes differ not only in the VS algorithm but also in other aspects. In particular, the control charts based on stepwise regression assume that the number of variables that can be potentially OC is fixed a priori; this condition has been relaxed for the LASSO- and LAR-based schemes (LEWMA and LAR-EWMA, hereafter). Indeed, these control charts assume that any, proper or improper, subset of the monitored variables can potentially shift. Further, LEWMA and LAR-EWMA are based on two slightly different control statistics. In addition, LAR-EWMA is developed not only for testing the status of the process mean but also for detecting an increase in the total variability.

To provide some guidelines on how to choose between different VS-based multivariate control charts and give some suggestions for further research, we here compare and discuss some VS-based control charts recently proposed in the SPC literature. For a more objective comparison, we use for all the investigated control charts the general regression model introduced in Capizzi and Masarotto (2011) for the LAR-EWMA. Indeed, this more general regression framework allows to handle a wide a variety of multivariate scenarios not only involving shifts in the component of a multivariate mean vector but also those related to changes in a profile or in a multistage process.

The paper is organized as follows. Section 2 briefly describes the procedures based on the variable selection algorithms. Section 3 presents the main results concerning comparisons, in terms of average run length (ARL), between some control schemes based on different VS-based algorithms. Details on the multivariate OC scenarios, discussed in the comparisons, are given in the Appendix. Concluding remarks are given in Sect. 4.

## 2 Statistical Monitoring Based on Variable Selection Algorithms

### 2.1 Generalities

Assume that at each time  $t$ ,  $t = 1, 2, \dots$ , independent observations on  $y_t$ , an  $n \times 1$  vector of quality characteristics, are available, and consider the following Gaussian change-point model

$$y_t \sim \begin{cases} N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}) & \text{if } t < \tau \text{ (in-control)} \\ N_n(\boldsymbol{\mu} + \boldsymbol{\delta}, \boldsymbol{\Omega}) & \text{if } t \geq \tau \text{ (out-of-control)} \end{cases} \quad (1)$$

that is, at  $\tau$ , an unknown instant of time, the mean vector and the covariance matrix shift leading the process to an OC state. Further, we suppose that the IC mean vector  $\boldsymbol{\mu}$  and the IC covariance matrix  $\boldsymbol{\Sigma}$  are known.

Concerning the OC mean vector, we assume that, at least approximately, the mean shift  $\boldsymbol{\delta}$  takes the form

$$\boldsymbol{\delta} = \mathbf{F}\boldsymbol{\beta}, \quad (2)$$

where  $\boldsymbol{\beta}$  is a  $p \times 1$  vector of unknown parameters and  $\mathbf{F}$  a suitable  $n \times p$  matrix of known constants. Thus, the mean vector may shift along any vector in the subspace spanned by the columns of  $\mathbf{F}$ , allowing for a multitude of potential shift directions. As shown in Capizzi and Masarotto (2011), formulation (2) is sufficiently flexible to encompass a wide variety of change-point scenarios. Further, suppose there is a practical interest only in detecting an increase in the total dispersion and assume that  $\boldsymbol{\Omega} - \boldsymbol{\Sigma}$  is a positive definite matrix.

Suppose process observations are accumulated in the following MEWMA

$$z_t = (1 - \lambda)z_{t-1} + \lambda(y_t - \boldsymbol{\mu}) \quad (3)$$

with  $z_0 = \mathbf{0}_n$ ,  $0 < \lambda \leq 1$ . Assuming the following (approximated) linear model

$$z_t = \mathbf{F}\boldsymbol{\beta} + \mathbf{a}_t,$$

with  $\mathbf{a}_t \sim N_n(\mathbf{0}_n, \lambda/(1 - \lambda)\boldsymbol{\Sigma})$ , the stability of the process mean can be checked by testing the hypothesis system

$$\begin{cases} H_0 : \boldsymbol{\beta} = \mathbf{0}_p, \\ H_1 : \boldsymbol{\beta} \neq \mathbf{0}_p. \end{cases} \quad (4)$$

Unfortunately, the standard test, described in any regression textbook, for the hypothesis system (4) can show a very low sensitivity when only a few components of  $\boldsymbol{\beta}$  effectively shift, and a much more efficient approach should consider alternative hypothesis systems on reduced subsets of the parameters.

A promising solution consists of using a suitable VS algorithm for determining subsets, having different sizes, of suspected variables, i.e., subsets of columns of  $F$  corresponding to nonzero coefficients. In particular, for  $k = 1, \dots, p$ , denote with  $J_k = \{j_{k,1}, \dots, j_{k,k}\}$  the indices of the selected predictors. Since the set of coefficients  $\{\beta_{j_{k,1}}, \dots, \beta_{j_{k,k}}\}$  correspond to a plausible subset of possible out-of-control parameters, the VS-based control statistics, for  $k = 1, \dots, p$ , test the following hypothesis systems:

$$\begin{cases} H'_0: \beta_j = 0 \text{ for } j = 1, \dots, p, \\ H'_{1,k}: \beta_j \neq 0 \text{ if } j \in J_k \text{ and } \beta_j = 0 \text{ if } j \notin J_k. \end{cases} \quad (5)$$

## 2.2 Three Different Approaches

Three distinct methods have been suggested for testing the hypothesis system (5).

In Wang and Jiang (2009) and Jiang et al. (2012), users are requested to choose in advance a suitable value for  $k$ . Then, for  $t = 1, 2, \dots$ , a standard forward search algorithm is used to select  $J_k$ , and an OC alarm is signaled when the following control statistic

$$S_{t,k} = \hat{\beta}'_{t,k} F' \Sigma^{-1} F \hat{\beta}_{t,k} \quad (6)$$

is greater than the control limit chosen for giving a desired IC performance. Here,  $\hat{\beta}_{t,k}$  denotes the GLS estimate of  $\beta$  obtained under  $H'_{1,k}$ , i.e., constraining to zero the coefficients of the predictors not in  $J_k$ .

In Zou and Qiu (2009),  $J_1, \dots, J_p$  are determined using the LASSO algorithm. Then, for  $k = 1, \dots, p$ , the authors suggest to compute the control statistic

$$V_{t,k} = \frac{(z'_t \Sigma^{-1} F \tilde{\beta}_{t,k})^2}{\tilde{\beta}'_{t,k} F' \Sigma^{-1} F \tilde{\beta}_{t,k}}, \quad (7)$$

where  $\tilde{\beta}_{t,k}$  denotes the LASSO estimator of  $\beta$  obtained under  $H'_{1,k}$ . An OC alarm is given when the overall control statistic

$$W_t = \max_{k=1, \dots, p} \frac{V_{t,k} - E[V_{t,k}]}{\sqrt{\text{Var}[V_{t,k}]}} \quad (8)$$

is greater than a suitable control limit. In (8), the mean and standard deviation of (7) are computed under the null hypothesis.

Alternatively, Capizzi and Masarotto (2011) suggest selecting  $J_1, \dots, J_k$  using the LAR algorithm and, for each  $k = 1, \dots, p$ , to compute the statistic  $S_{t,k}$ . Since it is important to detect not only changes in the process mean but also increases in

the dispersion, Capizzi and Masarotto (2011) also consider the additional alternative hypothesis

$$H'_{1,p+1} : \boldsymbol{\beta} = \mathbf{0}_p \text{ and } E\{(y_t - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (y_t - \boldsymbol{\mu})\} > n,$$

and the related one-sided EWMA statistic

$$S_{t,p+1} = \max \left( 1, (1 - \lambda) S_{t-1,p+1}^{(1)} + \lambda \frac{(y_t - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (y_t - \boldsymbol{\mu})}{n} \right), \quad (9)$$

with  $S_{0,p+1} = 1$ . Then, the LAR-based EWMA, for jointly monitoring the process mean and dispersion, is given by the aggregation of the  $p + 1$  statistics

$$M_t = \max_{k=1,\dots,p+1} \frac{S_{t,k} - E[S_{t,k}]}{\sqrt{\text{Var}[S_{t,k}]}} \quad (10)$$

where  $S_{t,k}$  is given by (6) for  $k = 1, \dots, p$ , and by (9) for  $k = p + 1$ . The combined control statistic (10) triggers an alarm when it exceeds a suitable control limit.

### 2.3 First Recommendations and Open Questions

As shown in Zou and Qiu (2009), Capizzi and Masarotto (2011), and Jiang et al. (2012), control charts like  $W_t$  and  $M_t$  offer a good protection against shifts occurring in one, two, ..., all components. Although the resulting scheme is not necessarily the best for detecting a shift occurring in a fixed number of components, it is usually close to the best. Conversely, control charts using a fixed value of  $k$ , such as those proposed by Wang and Jiang (2009) and Jiang et al. (2012), offer the best protection when shifts involve exactly  $k$  variables and unavoidably inferior protection when a shifts occur in a number of components different from the fixed value. Further, statistics such as  $W_t$  and  $M_t$  do not need an a priori choice of  $k$ . Thus, we suggest using an aggregated control statistic.

In addition, we strongly recommend including a control statistic, like  $S_{t,p+1}$ , designed for detecting a change in the dispersion. Indeed, joint monitoring of the process mean and dispersion is relevant *per se* but also provides some level of robustness against modeling errors and unforeseen behaviors. Further, as shown in the univariate case by Reynolds and Stoumbos (2005, 2006), the inclusion of a variance control statistic can be helpful for efficiently detecting large changes in the mean.

In the following, studying by simulation the ARL performance of VS-based control charts, we address the following additional issues: (1) Which variable selection algorithm should be used? (2) Which is better to use for monitoring, the elementary control statistic  $S_{t,k}$  or  $V_{t,k}$ ?

### 3 A Simulation Study

To address some of the issues discussed in the previous section, we compare five VS-based monitoring schemes. As recommended, all the schemes are based on a combination, similar to  $M_t$ , of  $p$  elementary control statistics used for detecting a mean shift and of the control statistic  $S_{t,p+1}$ , given by (9), for detecting increases in the total variation.

Details of the five control charts are given in Table 1. Note that, when a forward stepwise search is used, we have, for each  $k$ , that  $S_{t,k} = V_{t,k}$ . Thus, we present only one scheme for the forward VS algorithm. However, the control statistics  $V_{t,k}$ , given in (7), are here based on the LASSO- and LAR-based estimators of the vector  $\beta$ . Observe that in (6) and (7), at each stage  $k$ , the nonzero elements obtained via these three different VS algorithms are not necessarily the same.

Concerning the choice of the smoothing constant, as suggested in the literature (Lucas and Saccucci 1990; Prabhu and Runger 1997; Zou and Qiu 2009; Capizzi and Masarotto 2011; Jiang et al. 2012), a reasonable choice for normally distributed observations is between 0.1 and 0.3. The performance of the different VS-based schemes has been investigated for different values of  $\lambda$  and  $\tau$ . Because results are comparable for all the choices of these tuning constants, in the following results will be referred only to  $\lambda = 0.1$  and  $\tau = 1$ . The five VS-based control charts are compared in terms of out-of-control ARL evaluated using 500,000 Monte Carlo replications. The control limits, giving an in-control ARL equal to 500, have been computed using a stochastic approximation algorithm (Ruppert 1991; Polyak and Juditsky 1992). Within a reasonable number of iterations, the algorithm estimates the control limits with a given level of accuracy. Table 1 lists the estimates of the control limits for the five VS-based control charts. In addition, the mean and standard deviation of the elementary statistics  $S_{t,k}$  and  $V_{t,k}$ , for  $k = 1, \dots, p + 1$ , were computed by simulation.

Suitable choices of the matrix  $F$  lead to several change-point models, such as the “unstructured” scenario, when changes directly involve the components of the multivariate mean vector, and several “structured” scenarios, such those involving changes in a profile, that is, in the relationship between a response variable and

**Table 1** Five VS-based control charts

	FORWARD	LASSO/S	LASSO/V	LAR/S	LAR/V
VS algorithm	FORWARD	LASSO	LASSO	LAR	LAR
Elementary statistics	$S_{t,k}$	$S_{t,k}$	$V_{t,k}$	$S_{t,k}$	$V_{t,k}$
<i>Critical values</i>					
Unstructured	4.766477	4.578106	4.680325	4.722629	4.715809
Linear profile	5.033247	5.201960	5.011673	4.898843	5.061300
Cubic profile	4.877482	5.140879	4.889752	4.821646	4.870619
Nonparametric profile	5.268651	5.602346	5.546736	5.112253	5.156662
Multistage process	4.677853	4.603569	4.758084	4.726339	4.915792

one or more explanatory variables, and in a multistage process. Details for the components of the vector  $\boldsymbol{\beta}$  that are supposed to change are listed, for each change-point model, in the Appendix. In every case, several possible mean shifts are considered, including shifts in a single parameter, equal and different shifts in a pair of parameters, shifts of the same size in either even or odd components and shifts in variance. Here, we briefly describe the scenarios examined in the simulation study.

### 3.1 Unstructured

In this case  $p = n$ , the matrix  $\mathbf{F}$  reduces to the identity matrix  $\mathbf{F} = \mathbf{I}_n$  and the  $i$ -th element of  $\boldsymbol{\beta}$  directly points to a mean shift of the  $i$ -th quality characteristic, i.e.,  $\delta_i = \beta_i$ . Following the example in Zou and Qiu (2009), we consider  $p = n = 15$  and assume that the IC distribution is  $N_n(\mathbf{0}_n, \boldsymbol{\Sigma})$  with  $\boldsymbol{\Sigma} = (\sigma_{ij}) = (0.75^{|i-j|})$  for  $i, j = 1, 2, \dots, n$  and the OC distribution  $N_n(\boldsymbol{\beta}, \omega^2 \boldsymbol{\Sigma})$  with  $\omega > 1$ .

### 3.2 Linear and Cubic Profiles

Under this scenario, we assume that

$$y_{t,i} = \begin{cases} \epsilon_{t,i} & \text{if } t < \tau \\ \beta_1 + \beta_2 x_i + \dots + \beta_p x_i^{p-1} + \epsilon_{t,i} & \text{if } t \geq \tau \end{cases}$$

with  $x_i = (2i - n - 1)/(n - 1)$ , for  $i = 1, \dots, n$ . Here,  $\epsilon_{t,i}$  are independent, zero-mean, Gaussian random variables, with the IC and OC variance equal to one and  $\omega^2 > 1$ , respectively. Thus, in the described scenario,  $\mathbf{F} = (f_{i,j}) = (x_i^{j-1})$ ,  $\boldsymbol{\Sigma} = \mathbf{I}_n$  and  $\boldsymbol{\Omega} = \omega^2 \mathbf{I}_n$ . In particular, we consider linear ( $p = 2$ ) and cubic ( $p = 4$ ) profiles with  $n = 4$  and  $n = 8$  observations, respectively.

### 3.3 Nonparametric Profiles

To investigate the performance of the VS-based control chart for nonparametric monitoring of non-linear profiles, we use the same IC model considered by Zou et al. (2008),  $y_{t,i} = 1 - \exp(-x_i) + \epsilon_{t,i}$ , where  $x_i = (i - 0.5)/20$ ,  $i = 1, \dots, 20$ , and the following three OC models:

- I.  $y_{t,i} = 1 - \beta_1 \exp(-x_i^{\beta_2}) + \epsilon_{t,i}$ ;
- II.  $y_{t,i} = 1 - \exp(-x_{t,i}) + \beta_1 \cos(\beta_2 \pi (x_{t,i} - 0.5)) + \epsilon_{t,i}$ ;
- III.  $y_{t,i} = 1 - \exp(-x_{t,i} - \beta_1 \max(0, (x_{t,i} - \beta_2)/(1 - \beta_2))^2) + \epsilon_{t,i}$ .

Here,  $\epsilon_{t,i}$  are independent, zero-mean, Gaussian random variables, with the IC and OC variance equal to one and  $\omega^2 > 1$ , respectively. In this case, we set  $\mathbf{F}$  equal to the basis matrix of a cubic spline with four equispaced knots within the interval  $[0, 1]$ .

### 3.4 Multistage Processes

We consider an  $n$ -state process representable by the linear state-space model

$$\begin{cases} y_{t,i} = \mu_i + c_i x_{t,i} + v_{t,i} \\ x_{t,i} = d_i x_{t,i-1} + \beta_i I_{\{t \geq \tau\}} + w_{t,i} \end{cases} \quad (i = 1, \dots, n)$$

where  $v_{t,i}$  and  $w_{t,i}$  are independent normal random variables with zero mean. The  $n$  elements of the  $\boldsymbol{\beta} = (\beta_i)$  vector define the magnitude of the shifts and the stages at which the shifts occur. It is easy to show that this model is a particular case of (1). See Capizzi and Masarotto (2011) for the details and, in particular, for the structure of the  $\mathbf{F}$  and  $\boldsymbol{\Sigma}$  matrices. In the simulation, we fix the number of stages to  $n = 10$  and investigate the performance for different shift locations, occurring in one, two, five and all stages assuming that  $\mu_i = 0$ ,  $c_i = d_i = \text{var}(w_{t,i}) = \text{var}(v_{t,i}) = 1$  for every  $i$ .

### 3.5 Results

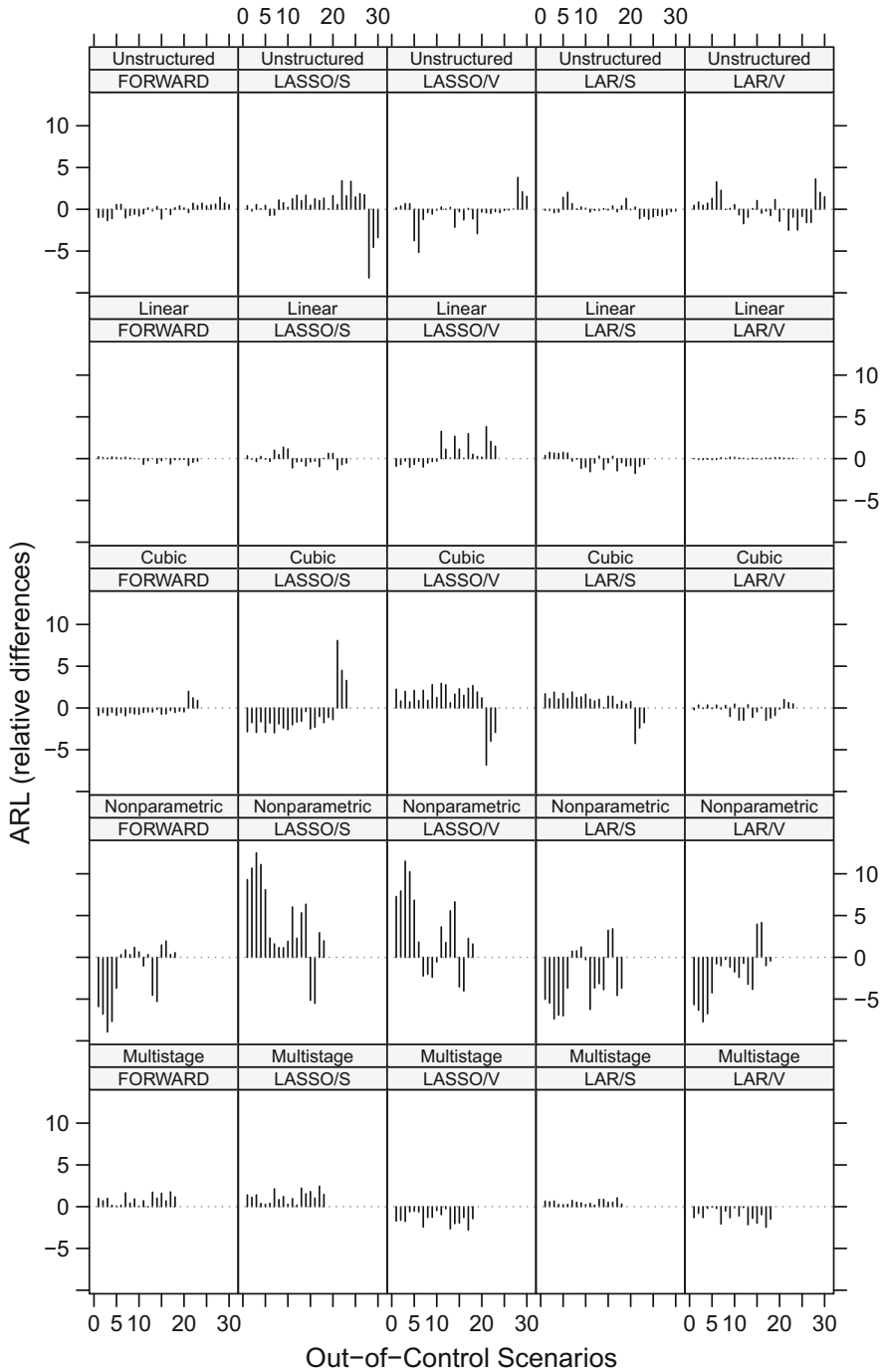
Results are summarized in Fig. 1, which shows the following percent relative differences

$$100 \cdot \frac{\text{ARL}_{rs} - \text{MARL}_r}{\text{MARL}_r}, \quad s = 1, \dots, 5, \quad (11)$$

i.e., the percent relative differences between  $\text{ARL}_{rs}$ , the OC ARL of  $s$ -th control chart in the  $r$ -th OC scenario, and  $\text{MARL}_r$ , the mean of the five out-of-control ARL values, one for each control chart, obtained for the  $r$ -th OC scenario. Observe that the number of the OC scenarios is different for the different cases. In particular,  $r = 1, \dots, 18$  for the case of nonparametric and multistage process monitoring,  $r = 1, \dots, 23$  for the monitoring of linear and cubic profiles and  $r = 1, \dots, 30$  in the unstructured case (see the Appendix for a detailed description of each OC scenario).

A negative (positive) value of (11) can be interpreted as a quicker reaction (slower) reaction of the  $s$ -th control chart to the  $r$ -th OC situation, when compared to the other VS-based control charts.





**Fig. 1** Relative ARL differences of five VS-based control charts for several OC scenarios (see the Appendix for labels in the  $x$  axis)

Results show that, independently from the multivariate control charts, the forward- and the LAR-based schemes show a similar behavior that seems to be quite stable throughout all the different practical contexts. However, while LASSO shows a substantially negligible advantage for some out-of-control situations, it can also show a relatively large degradation for other applications, such as nonparametric profile monitoring. Further, monitoring schemes based on the same VS algorithm but on different elementary control statistics, i.e., on  $S_{t,k}$  or  $V_{t,k}$ , offer essentially the same performance.

## 4 Conclusions

In this paper, we compared the performance of multivariate control charts based on three different variable selection procedures. In particular, the compared multivariate control charts have been implemented for detecting many out-of-control conditions, even involving increases in process variation, for several MSPC frameworks. Results show that whereas control charts consisting of the aggregation of several control statistics, such as those proposed by Zou and Qiu (2009) and Capizzi and Masarotto (2011), behave quite similarly for different OC applications, suggestions can be given to practitioners concerning the particular variable selection procedure to use. As discussed before, the LASSO-based control charts can show an unsatisfactory performance in detecting some particular OC situations. However, the forward and LAR-based procedures can be considered substantially equivalent in terms of OC ARL performance for the investigated practical applications. Thus, from a practical point of view, multivariate control charts based on forward selection could be more appealing to users since these charts are more intuitive and simpler to implement.

## Appendix

In the following, we provide details on the OC scenarios listed in the  $x$  axes of Fig. 1. When only variables or stages with an even (odd) index are subject to a shift of size  $\delta$ , the OC scenario is indicated with either Even $[\delta]$  or Odd $[\delta]$ .

1. *Unstructured model*: 1 = ( $\beta_1 = 0.5$ ), 2 = ( $\beta_1 = 1$ ), 3 = ( $\beta_3 = 0.5$ ), 4 = ( $\beta_3 = 1$ ), 5 = ( $\beta_1 = 0.5, \beta_2 = 0.25$ ), 6 = ( $\beta_1 = 0.5, \beta_2 = 0.5$ ), 7 = ( $\beta_1 = 0.5, \beta_2 = 0.75$ ), 8 = ( $\beta_1 = 0.5, \beta_3 = 0.25$ ) 9 = ( $\beta_1 = 0.5, \beta_3 = 0.5$ ), 10 = ( $\beta_1 = 0.5, \beta_3 = 0.75$ ), 11 = ( $\beta_3 = 0.5, \beta_8 = 0.25$ ), 12 = ( $\beta_3 = 0.5, \beta_8 = 0.5$ ), 13 = ( $\beta_3 = 0.5, \beta_8 = 0.75$ ), 14 = ( $\beta_1 = 0.5, \beta_2 = 0.25, \beta_3 = 0.25$ ), 15 = ( $\beta_1 = 0.25, \beta_2 = 0.25, \beta_3 = 0.5$ ), 16 = ( $\beta_2 = 0.5, \beta_3 = 0.25, \beta_8 = 0.25$ ), 17 = ( $\beta_2 = 0.25, \beta_3 = 0.25, \beta_8 = 0.5$ ), 18 = ( $\beta_7 = 0.5, \beta_8 = 0.25, \beta_9 = 0.5$ ), 19 = ( $\beta_7 = 0.25, \beta_8 = 0.75, \beta_9 = 0.5$ ), 20 = ( $\beta_6 = 0.5, \beta_8 = 0.25, \beta_{10} = 0.5$ ), 21 = ( $\beta_6 = 0.25, \beta_8 = 0.75, \beta_{10} = 0.5$ ),

- 22 = (Even[0.25]), 23 = (Even[0.5]), 24 = (Odd[0.25]), 25 = (Even[0.5]),  
 26 = (Even[0.5], Odd[0.25]), 27 = (Odd[0.5], Even[0.25]), 28 = ( $\omega = 1.2$ ),  
 29 = ( $\omega = 1.5$ ), 30 = ( $\omega = 2$ ).
2. *Linear profiles:* 1 = ( $\beta_1 = 0.1$ ), 2 = ( $\beta_1 = 0.3$ ), 3 = ( $\beta_2 = 0.2$ ), 4 = ( $\beta_2 = 0.5$ ), 5 = ( $\beta_3 = 0.2$ ), 6 = ( $\beta_3 = 0.5$ ), 7 = ( $\beta_1 = 0.1$ ,  $\beta_2 = 0.1$ ), 8 = ( $\beta_1 = 0.4$ ,  $\beta_2 = 0.2$ ) 9 = ( $\beta_1 = 0.4$ ,  $\beta_2 = 0.6$ ), 10 = ( $\beta_1 = 0.6$ ,  $\beta_2 = 0.6$ ), 11 = ( $\beta_1 = 0.1$ ,  $\omega = 1.2$ ), 12 = ( $\beta_1 = 0.3$ ,  $\omega = 1.2$ ), 13 = ( $\beta_1 = 1$ ,  $\omega = 1.2$ ), 14 = ( $\beta_2 = 0.1$ ,  $\omega = 1.2$ ), 15 = ( $\beta_2 = 0.4$ ,  $\omega = 1.2$ ), 16 = ( $\beta_1 = 1.2$ ,  $\omega = 1.2$ ), 17 = ( $\beta_1 = 0.1$ ,  $\beta_2 = 0.1$ ,  $\omega = 1.2$ ), 18 = ( $\beta_1 = 0.4$ ,  $\beta_2 = 0.2$ ,  $\omega = 1.2$ ), 19 = ( $\beta_1 = 0.4$ ,  $\beta_2 = 0.6$ ,  $\omega = 1.2$ ), 20 = ( $\beta_1 = 0.6$ ,  $\beta_2 = 0.6$ ,  $\omega = 1.2$ ), 21 = ( $\omega = 1.2$ ), 22 = ( $\omega = 1.5$ ), 23 = ( $\omega = 2$ ).
3. *Cubic profiles:* 1 = ( $\beta_1 = 0.1$ ), 2 = ( $\beta_1 = 0.3$ ), 3 = ( $\beta_1 = 1$ ), 4 = ( $\beta_2 = 0.2$ ), 5 = ( $\beta_2 = 0.4$ ), 6 = ( $\beta_2 = 1.2$ ), 7 = ( $\beta_4 = 0.2$ ), 8 = ( $\beta_4 = 0.5$ ), 9 = ( $\beta_1 = 0.1$ ,  $\beta_2 = 0.1$ ), 10 = ( $\beta_1 = 0.1$ ,  $\beta_3 = 0.2$ ) 11 = ( $\beta_1 = 0.1$ ,  $\beta_4 = 0.2$ ), 12 = ( $\beta_2 = 0.2$ ,  $\beta_3 = 0.2$ ), 13 = ( $\beta_2 = 0.4$ ,  $\beta_4 = 0.2$ ), 14 = ( $\beta_3 = 0.5$ ,  $\beta_4 = 0.5$ ), 15 = ( $\beta_1 = 0.1$ ,  $\beta_2 = 0.1$ ,  $\beta_3 = 0.1$ ), 16 = ( $\beta_1 = 0.1$ ,  $\beta_3 = 0.2$ ,  $\beta_4 = 0.1$ ), 17 = ( $\beta_2 = 0.1$ ,  $\beta_3 = 0.3$ ,  $\beta_4 = 0.2$ ), 18 = ( $\beta_1 = 0.1$ ,  $\beta_2 = 0.1$ ,  $\beta_3 = 0.1$ ,  $\beta_4 = 0.1$ ), 19 = ( $\beta_1 = 0.1$ ,  $\beta_2 = 0.2$ ,  $\beta_3 = 0.1$ ,  $\beta_4 = 0.2$ ), 20 = ( $\beta_1 = 0.2$ ,  $\beta_2 = 0.1$ ,  $\beta_3 = 0.2$ ,  $\beta_4 = 0.1$ ), 21 = ( $\omega = 1.2$ ), 22 = ( $\omega = 1.5$ ), 23 = ( $\omega = 2$ ).
4. *Non parametric profiles.* The following OC scenarios are referred to possible shifts in the regression coefficients of models I, II and III. 1 = (I,  $\beta_1 = 1.00$ ,  $\beta_2 = 1.30$ ), 2 = (I,  $\beta_1 = 1.00$ ,  $\beta_2 = 1.50$ ), 3 = (I,  $\beta_1 = 1.10$ ,  $\beta_2 = 1.00$ ) 4 = (I,  $\beta_1 = 1.30$ ,  $\beta_2 = 1.00$ ), 5 = (I,  $\beta_1 = 1.20$ ,  $\beta_2 = 1.00$ ,  $\omega = 1.10$ ), 6 = (I,  $\beta_1 = 1.00$ ,  $\beta_2 = 1.20$ ,  $\omega = 1.30$ ), 7 = (II,  $\beta_1 = 0.10$ ,  $\beta_2 = 3.00$ ), 8 = (II,  $\beta_1 = 0.30$ ,  $\beta_2 = 3.00$ ), 9 = (II,  $\beta_1 = 0.10$ ,  $\beta_2 = 2.00$ ) 10 = (II,  $\beta_1 = 0.30$ ,  $\beta_2 = 2.00$ ), 11 = (II,  $\beta_1 = 0.20$ ,  $\beta_2 = 4.00$ ,  $\omega = 1.10$ ), 12 = (II,  $\beta_1 = 0.20$ ,  $\beta_2 = 4.00$ ,  $\omega = 1.30$ ) 13 = (III,  $\beta_1 = 2.00$ ,  $\beta_2 = 0.90$ ), 14 = (III,  $\beta_1 = 4.00$ ,  $\beta_2 = 0.90$ ), 15 = (III,  $\beta_1 = 2.00$ ,  $\beta_2 = 0.75$ ), 16 = (III,  $\beta_1 = 4.00$ ,  $\beta_2 = 0.75$ ), 17 = (III,  $\beta_1 = 2.00$ ,  $\beta_2 = 0.90$ ,  $\omega = 1.20$ ), 18 = (III,  $\beta_1 = 4.00$ ,  $\beta_2 = 0.75$ ,  $\omega = 1.20$ ).
5. *Multistage process.* In the following,  $\beta_j$  indicates the shift of size  $\beta$ , occurring at the  $j$ -th stage,  $j = 1, \dots, 20$ . 1 = ( $\beta_1 = 0.75$ ), 2 = ( $\beta_5 = 0.75$ ), 3 = ( $\beta_{10} = 0.75$ ), 4 = ( $\beta_1 = 1.5$ ), 5 = ( $\beta_5 = 1.5$ ), 6 = ( $\beta_{10} = 1.5$ ), 7 = ( $\beta_2 = 0.6$ ,  $\beta_8 = 0.6$ ), 8 = ( $\beta_4 = 0.6$ ,  $\beta_5 = 0.6$ ), 9 = ( $\beta_2 = 1.2$ ,  $\beta_8 = 1.2$ ), 10 = ( $\beta_4 = 1.2$ ,  $\beta_5 = 1.2$ ), 11 = ( $\beta_2 = 1.8$ ,  $\beta_8 = 1.8$ ), 12 = ( $\beta_4 = 1.8$ ,  $\beta_5 = 1.8$ ), 13 = ( $\beta_1 = \beta_5 = \beta_{10} = 0.2$ ,  $\beta_3 = \beta_7 = 0.4$ ), 14 = ( $\beta_3 = \beta_5 = \beta_7 = 0.2$ ,  $\beta_4 = \beta_6 = 0.4$ ), 15 = ( $\beta_1 = \beta_5 = \beta_{10} = 0.6$ ,  $\beta_3 = \beta_7 = 0.4$ ), 16 = ( $\beta_3 = \beta_5 = \beta_7 = 0.6$ ,  $\beta_4 = \beta_6 = 0.4$ ), 17 = (Even[0.1], Odd[0.2]), 18 = (Even[0.5], Odd[0.25]).

## References

- Bersimis, S., Psarakis, S., & Panaretos, J. (2007). Multivariate statistical process control charts: an overview. *Quality and Reliability Engineering International*, 23, 517–543.
- Capizzi, G., & Masarotto, G. (2011). A least angle regression control chart for multidimensional data. *Technometrics*, 53, 285–296.
- Efron, E., Hastie, T., Johnstone, I., & Tibshirani, R. (2004). Least angle regression. *The Annals of Statistics*, 32, 407–499.
- Jiang, W., Wang, K., & Tsung, F. (2012). A variable-selection-based multivariate EWMA chart for process monitoring and diagnosis. *Journal of Quality Technology*, 44, 209–230.
- Lucas, J. M., & Saccucci, M. S. (1990). Exponentially weighted moving average control schemes: properties and enhancements. *Technometrics*, 32, 1–29.
- Polyak, B. T., & Juditsky, A. B. (1992). Acceleration of stochastic approximation by averaging. *SIAM Journal of Control Optimization*, 30, 838–855.
- Prabhu, S. S., & Runger, G. C. (1997). Designing a multivariate EWMA control chart. *Journal of Quality Technology*, 29, 8–15.
- Reynolds, M. R., & Stoumbos, Z. G. (2005). Should exponentially weighted moving average and cumulative sum charts be used with Shewhart limits? *Technometrics*, 47(4), 409–424.
- Reynolds, M. R., & Stoumbos, Z. G. (2006). Comparisons of some exponentially weighted moving average control charts for monitoring the process mean and variance. *Technometrics*, 48(4), 550–567.
- Ruppert, D. (1991). Stochastic approximation. In B. K. Ghosh, & P. K. Sen (Eds.), *Handbook of sequential analysis* (pp. 503–529). New York, NY: Marcel Dekker.
- Tibshirani, R. J. (1996). Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society, Series B*, 58, 267–288.
- Wang, K., & Jiang, W. (2009). High-dimensional process monitoring and fault isolation via variable selection. *Journal of Quality Technology*, 41, 247–258.
- Zou, C., Ning, X., & Tsung, F. (2010). Lasso-based multivariate linear profile monitoring. *Annals of Operation Research*, 192, 3–19.
- Zou, C., & Qiu, P. (2009). Multivariate statistical process control using LASSO. *Journal of American Statistical Association*, 104, 1586–1596.
- Zou, C., Tsung, F., & Wang, Z. (2008). Monitoring profiles based on nonparametric regression methods. *Technometrics*, 50, 512–526.