# A Multivariate Change point Model for Change in Mean Vector and/or Covariance Structure: Detection of Isolated Systolic Hypertension (ISH).

K.D. Zamba, Ph.D.[1]

[1] College of Public Health, Department of Biostatistics, The University of Iowa, IA;
[1] The Holden Comprehensive Cancer Center, The University of Iowa, IA.

## Abstract

The motivation of the current research manuscript is to provide practitioners with a multivariate analysis tool able to detect change in the mean vector and/or covariance matrix, as well as the epoch of a change, in an independent sequence of multivariate observations. The work explores the multivariate change point model through generalized likelihood ratio statistics applied sequentially, and adapted to repeated use. We sought an analytical result for the exact moments of the generalized likelihood statistic. The benefit flowing from this sequential adaptation is to be able to detect the epoch and the source of potential changes in independent multivariate readings. Possible areas of application are: ambulatory monitoring, disease monitoring, syndromic surveillance, and sequential dynamic control.

## Keywords

Likelihood ratio, Change point, Sequential analysis.

# Introduction and Objective

Multivariate statistical process control carries out ongoing checks to ensure that a process is in-control. These checks are traditionally done by the $T^2$, multivariate cusum, and multivariate exponentially weighted moving average (MEWMA) control charts. These traditional tools rely on known or assumed known in-control true parameters, and use the assumed true values to set the control limits. The reality however is that true parameter values are seldom, if ever, known exactly; rather they are commonly estimated from a Phase I sample. It is increasingly being recognized that the Phase I study needs to involve large samples if the random errors in the parameter estimates are to provide run behavior matching that of known-parameter situation. Apart from the general undesirability of large and therefore expensive studies preliminary to actual charting, some control settings have a paucity of relevant data to estimate the process parameters. Zamba & Hawkins (2006) outlined the advantage of using the unknown parameter likelihood ratio tests in multivariate control. Although their manuscript applies specifically to mean-only change point models, the idea can be generalized to the mean and/or covariance change point models.

In multivariate control, change in a process mean vector can be masked by unsuspected change in covariance structure or by a sudden change in the correlation between two quality characteristics. For a deeper insight on this aspect of the relationship between quality characteristics and their repercussion on the performance of a chart, we refer the reader to Hotelling (1947), Mason & Young (2002). The objectives of multivariate statistical process control do not differ much from those of univariate. They include providing a signal that the process is out of control; an estimate of when it went out of control; and a diagnosis of the way in which it went out of control – for example, whether some or all of the components of the mean vector have shifted, whether the covariance matrix is perturbed and some quality characteristics became counter-correlated, or whether both the mean vector and covariance matrix undergo some change. The most widely setting for multivariate SPC is the multivariate normal. This setting, which we will use, assumes that while in control, the readings follow independent common multivariate normal distributions with some mean

vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. This article stresses the possibility that the mean vector and/or covariance matrix may have changed from ($\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$) to ($\boldsymbol{\mu}_1$, $\boldsymbol{\Sigma}_1$). We further focus on an out of control state due to *persistent* or *sustained* causes, by which we mean a process has left the state of control, remains out of control, or even goes further from control, until some corrective action is taken. The setting in which an out of control state is *transient* or *isolated*, by which we mean that the system goes out of control but then returns to control even in the absence of any intervention; is not the focus of this paper.

# Existing Work

The original and best known work in multivariate control chart is that of Hotelling (1947) which is a direct multivariate equivalent of the Shewhart $\bar{X}$ chart. Conventional univariate methods that are best suited for persistent small changes, such as cumulative sum (cusum) chart and the exponentially weighted moving average (EWMA), are by and large limited in the multivariate arena. Multivariate versions of the cusum tend to be somewhat specialized. See for example Healy (1987), Crosier (1988), Pignatiello and Runger (1990), Hawkins and Olwell (1998). It may be fair to say that there is no multivariate cusum *per se* other than the proposal by Crosier (1988); rather, there are proposals for a collection of univariate cusums. For identifying shifts in mean for a known in-control value $\boldsymbol{\mu}$ and a specified shift of size $\boldsymbol{\delta}$, the optimal diagnostic was shown by Healy (1987) to be the univariate cusum of the scalar $(\boldsymbol{X} - \boldsymbol{\mu} - \boldsymbol{\delta}/2)'\boldsymbol{\Sigma}^{-1}\boldsymbol{\delta}$. This result though is usually of no particular use since we do not know in what direction the mean will shift and, unlike the scalar case where mis-designing the cusum leads to no more than the loss of some of the performance possible with more accurate tuning, if the direction specified is wrong, the performance of the cusum may be arbitrarily bad. There have been proposals for cusum charts of multivariate data with a limited number of likely modes of failure (Hawkins and Olwell 1998), but except in special circumstances these generally lack the simplicity and robust good performance of their univariate counterparts. The multivariate version of the EWMA, due to Lowry et al.(1992), is sensitive to small persistent mean shifts in any direction. It can also be used to

detect larger shifts by using larger values of its tuning constant.

The cusum and the MEWMA are primarily intended to identify persistent changes in the mean vector. Changes in the covariance matrix have traditionally received less attention. Two recent proposals (Huwang et al 2006, Hawkins and Maboudou-Tchao 2008) have extended the MEWMA to the monitoring of the covariance matrix. In these proposals, the process data vectors $\boldsymbol{X}_n$ are first multi-standardized to $\boldsymbol{Y}_n = \boldsymbol{\Sigma}^{-1/2}(\boldsymbol{X}_n - \boldsymbol{\mu})$, which are $N(0, \boldsymbol{I})$ while the process is in control. A running estimated covariance matrix is then defined by the recursion

$$\boldsymbol{S}_n = (1-\gamma)\boldsymbol{S}_{n-1} + \gamma \boldsymbol{Y}_n \boldsymbol{Y}'_n$$

with the initialization $\boldsymbol{S}_0 = 0$ (Huwang et al, 2007) or $\boldsymbol{S}_0 = \boldsymbol{I}$ (Hawkins and Maboudou-Chao, 2007).

These proposals assume that the in-control process parameters are known exactly. The $T^2$chart for estimated parameters has been recently studied by Champ et al.(2005) and designed for rational groups of size 3 or more. Their performance analysis attests to the undesirable effect of estimation on the run behavior of a chart – while standard statistical theory allows one to set control limits with any desired probability of exceedance, the use of the same parameter estimates in different rational groups' $T^2$ creates serial dependence that distorts the run length distribution.

A conceptually different approach to SPC in the setting of unknown in-control parameters is the unknown-parameter change point formulation. Hawkins Qiu and Kang (2003), Hawkins and Zamba (2005,a,b), Zamba and Hawkins(2006) outlined reasons for using sequential change point methodologies in a process control where the location parameters are unknown or not fully known (Hawkins Qiu and Kang, Zamba and Hawkins), where univariate scale and/or location parameters are not fully known(Hawkins and Zamba a,b) . The reasons can be summarized as follow:

- Avoid the problem of dependence on assumed known parameter values
- Being able to chart startup processes and low volume productions

- Being able to monitor and learn simultaneously
- Being able to control the run behavior

In the following, we define the multivariate unknown parameter change point model, their sequential and dynamic use.

# The Unknown-Parameter Change Point Model

The working model

$$\boldsymbol{X}_i \sim \begin{cases} N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}) & \text{if } i \leq \tau \\ N_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1) & \text{if } i > \tau \end{cases} \tag{1}$$

defines the change point formulation in a $p$-variate normal case. The parameter $\tau$ is called the change point (if one exists), and is unknown. The remaining parameters $\boldsymbol{\mu}$, $\boldsymbol{\mu}_1$, $\boldsymbol{\Sigma}$, $\boldsymbol{\Sigma}_1$ are the in- and out-of-control mean vectors and covariance matrices. While there are change point formulations with some of the parameters assumed known, we will be focussing on the unknown-parameter change point problem in which none of the parameters is assumed known *a priori.* We refer the reader to Zamba & Hawkins (2006) for a literature review on the static applications of multivariate change point methods. There are various specializations of the general model (1). The assumption that $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_1$ leads to a model where the change is in the mean only, a model discussed in Zamba and Hawkins (2006). Sketching the development, if in a series of $n$ observations a change point were known *a priori* to be at $\tau = k$, then $T_k^2$, the Hotelling-$T^2$ between pre- and post-$k$ data, would be the likelihood ratio test statistic for testing a change between the pre-$k$ and post-$k$ segment means. In case the change point is not known ahead of time, the maximum over all possible split points

$$T_{\hat{\tau}}^2 = \max T_k^2, \quad k = 1, ..., n-1$$

is the generalized likelihood ratio test statistic for change in mean vector. The maximizing index $\hat{\tau}$ is the maximum likelihood estimate (MLE) of the change point.

When the generalized likelihood ratio test statistic is used sequentially to assess a change in mean vectors of two segments, as a new data vector $\boldsymbol{X}_n$ accrues, the algorithm follows:

- Compute $T^2_{\max,n}$ the maximized split statistic for the entire sequence of readings $\boldsymbol{X}_1, \boldsymbol{X}_2, ...\boldsymbol{X}_n$ to date.
- If $T^2_{\max,n}$ exceeds some cutoff $h_{n,p,\alpha}$, conclude that there has been a shift. We estimate the time of occurrence of the shift by the maximizing index $k$, and for follow-up diagnosis conduct a two-sample study of the pre-$k$ and post-$k$ series to identify the change.
- If $T^2_{\max,n} < h_{n,p,\alpha}$, conclude that there is not enough evidence of a change, and let the process continue without interruption.

The thresholds $h_{n,p,\alpha}$ are based on a desirable property of the sequence of length $n$ to give a constant probability of false alarm for each $n$ (Hawkins, Qiu and Kang (2003), Margavio *et al.* (1995)). If this probability were a constant $\alpha$ say, then the sequence must satisfy the equation

$$P[T^2_{\max,n} > h_{n,p,\alpha} \mid T^2_{\max,j} \leq h_{j,p,\alpha}; \quad j < n] = \alpha. \tag{2}$$

## The Change point Model for Change in Mean and/or Covariances: The Adaptation

The adaptation to a more general setting where mean and covariance matrix may undergo a sudden change is outlined as follow. We assume the existence of a time $\tau$ when the mean vector and/or the covariance matrix change. The change point model that encapsulates this scenario is the general setting where

$$\boldsymbol{X}_i \sim \begin{cases} N_p(\boldsymbol{\mu}\ ,\ \boldsymbol{\Sigma}) & \text{if} \quad i \leq \tau, \\ N_p(\boldsymbol{\mu}_1\ ,\ \boldsymbol{\Sigma}_1) & \text{if} \quad i > \tau. \end{cases} \tag{3}$$

Under the assumption of no change, this model (3) can be summarized in terms of joint hypothesis as follow

$$H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_1 \;\; ; \;\; \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_1 \quad .$$

This hypothesis can be split into the combination of two composite hypotheses; the first of which is

$$H_{0,1} : \boldsymbol{\mu} = \boldsymbol{\mu}_1 \;\; | \;\; \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_1 \; ,$$

and the second

$$H_{0,2} \; : \; \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_1 \quad .$$

Note that in $H_{0,1}$, we have knowledge that the pre and post-shift covariance matrices are the same. This setting yields the mean-only change point model as discussed in Zamba & Hawkins (2006). $H_{0,2}$ is a covariance-only change point model. Considering the unrestricted parameter space defined under $H_0$, and the restricted parameter space defined under $H_{0,1}$, if an observation falls within the space defined by $H_0$, it should primarily fall into the restricted space defined under $H_{0,1}$ given that it falls in the space defined under $H_{0,2}$.

Write $\bar{X}_{i,j}$ and $\boldsymbol{S}_{i,j}$ for the mean vector and the maximum-likelihood sample covariance matrix for the vectors $\boldsymbol{X}_{i+1}, \boldsymbol{X}_{i+2}, ...\boldsymbol{X}_j$ - that is

$$\bar{X}_{i,j} = \frac{1}{(j-i)} \sum_{k=i+1}^{j} \boldsymbol{X}_k, \quad \boldsymbol{S}_{i,j} = \sum_{k=i+1}^{j} (\boldsymbol{X}_k - \bar{X}_{i,j})(\boldsymbol{X}_k - \bar{X}_{i,j})'/(j-i).$$

By Lemma 10.3.1, Anderson(1984), the likelihood ratio test statistic for testing $H_0$ assuming a split point at observation $k$ is

$$\Lambda_{k,n} \;\; \propto \;\; \frac{\mid \boldsymbol{S}_{0,k} \mid^{\frac{k-1}{2}} \times \mid \boldsymbol{S}_{k,n} \mid^{\frac{n-k-1}{2}}}{\mid \boldsymbol{S}_{0,n} \mid^{\frac{n-1}{2}}}, \tag{4}$$

and the familiar doubled negative log likelihood ratio is

$$-2 \log \Lambda_{k,n} = (n-1) \log |\boldsymbol{S}_{0,n}| - (k-1) \log |\boldsymbol{S}_{0,k}| - (n-k-1) \log |\boldsymbol{S}_{k,n}|.$$

### A Bartlett Correction

The generalized likelihood ratio test has an asymptotic chi-squared distribution with $p(p+3)/2$ degrees of freedom. This asymptotic distribution would be relevant if both $k$ and $n-k$ were large. However in SPC applications, fast response is important, leading to the conclusion that SPC procedures must be able to handle settings in which a segment or segments is short. For this reason, it is advisable to use the Bartlett correction of normalizing the GLR by its null expectation, a procedure that is known to dramatically improve the chi-squared approximation.

Under the in-control distribution with the sequence of length $n$ split at point $k$, the mean of the log likelihood ratio statistic is given by: $E(-2\log\Lambda_{k,n}) = g_{k,n}$, where

$$\begin{aligned} g_{k,n} = \quad & p(\log 2 + (n-1)\log(n-1) - (n-k-1)\log(n-k-1) - (k-1)\log(k-1)) \\ & + \sum_{j=1}^{p}[(n-1)\psi(\frac{n-j}{2}) - (k-1)\psi(\frac{k-j}{2}) - (n-k-1)\psi(\frac{n-k-j}{2})], \end{aligned} \tag{5}$$

with $\psi(z) = \frac{d}{dz}\log\Gamma(z)$, the digamma function.

The Bartlett correction then consists of using the test statistic

$$G_{k,n} = -2\log\Lambda_{k,n}(p(p+3))/g_{k,n} \tag{6}$$

which, for a fixed $k$, has a scaled approximate chi-squared distribution.

This result is outlined in the Appendix.

### Application – Sequential Change Point Procedure

Analogous with the earlier change point proposals, we suggest the following procedure for detection of change points in mean vector and/or covariance matrix.

- For each new observation $n$, compute $G_{k,n}$ for each $k$ in the feasible range $p+1, ..., n-p-1$.
- Calculate $G_{\max,n}$, the maximum of these $G_{k,n}$.

- If $G_{\max,n}$ exceeds a control limit $h_{n,p,\alpha}$, then signal a change. Diagnose the epoch of the change as the $k$ value yielding the maximum.

- Otherwise continue to the next observation vector.

The control limits $h_{n,p,\alpha}$ are chosen to fix the probability of a false alarm at each observation to $\alpha$. In other words, they are the solution to the equation:

$$P[G_{\max,n} > h_{n,p,\alpha} \mid G_{\max,j} \leq h_{j,p,\alpha}; j < n] = \alpha. \tag{7}$$

As an analytical solution to the equations seems intractable, they were estimated by simulation. This involved 5 million simulations to estimate the limits for series lengths of up to 150. We covered $\alpha$ values of 0.0005, 0.001, 0.002, 0.005 and 0.01 corresponding respectively to fixed ARL of 2000, 1000, 500, 200 and 100, and examined dimensions $p =\{2, 3, 4, 5, 10, 15, 20\}$.

In a purely mathematical sense, the unknown-parameter change point formulation can be run with $n$ as small as $2(p+1)$, which is the minimum sample size required for both $\boldsymbol{S}_{0,k}$ and $\boldsymbol{S}_{k,n}$ to have a chance of being non-singular. Most practitioners though would accumulate 'a few' observations beyond the mathematical lower limit before starting testing. Our simulations incorporate this belief by allowing for some number of 'learning' observations to accrue before formal monitoring starts. We set the number of these learning observations to 0, 10, 20, 30, 40 in addition to the $2(p+1)$ required for non-singularity. Tables of estimated control limits along with their standard error on our are available upon request. Table 1 is an extract from these more extensive tables. It shows the cut points for ARL = 500, assuming an initial no-test learning period totalling $2(p+1)+10$ process vectors: choices that we believe would suit many practitioners.

# Basic Performance study

A full-scale performance evaluation of our proposal is challenging in view of the large number of factors affecting performance. There are the immediate factors of the dimensionality and

the in-control ARL. Then there is the fact that the process mean can change in an arbitrary way, and that any subset of the $p(p+1)/2$ elements of the covariance matrix could change. Finally, there is the impact of the length of time the process runs in control before the shift – it is intuitively clear that a long in-control run leads to good estimates of the in-control parameters, and so should to some degree speed the response to an out-of-control situation. In order to explore the impact of some of these factors, we conducted a basic simulation study. Various runs used initial in-control series of length 10, 20, 30, and 40 observations. We then added shifts in mean and/or variance following the initial sequence, and applied our procedure. Since the generalized likelihood criterion is not affected by a full-rank linear transformation of the vectors $\boldsymbol{X}_j$; Anderson (1984), to explore arbitrary changes in the mean vector, it is sufficient to shift a single component of the mean. To study location shifts, we therefore added a shift $\delta$ to the first component of each data vector following the in-control sequence. Regarding the covariance matrix, we multiply one component of our covariance matrix by $1+\lambda$. For example, assuming the component is the first, for a $p \times p$ matrix $\Sigma_p = \begin{pmatrix} \sigma_{1,1} & \dots & \sigma_{1,p} \\ \sigma_{p,1} & ( \ \Sigma_{p-1} \ ) & \end{pmatrix}$, multiplying the first entry $\sigma_{1,1}$ by $1+\lambda$ to compute the determinant translates into multiplying the determinant by $(1+\lambda)\times \mid \Sigma_{p-1} \mid$. One can choose another entry and have a change in the determinant as a function of the cofactor of the entry selected. We chose the first component for simplicity. After these changes, we ran our chart looking for a signal. We recorded the $n$ at which the chart signaled a shift and the estimate of the change point. A false alarm occurs when a series signals before the end of its in-control readings. For $\delta = 0$, and $\lambda = 0$ all alarms are false, and are expected to occur at rate $\alpha$ giving the chart an ARL of $1/\alpha$. In our simulation, the charts were all tuned to in control ARL of 500, corresponding to a false alarm rate $\alpha = 0.002$. We ran 10,000 replications of length $n = 2000$ each, using dimensions 2, 5 and 10. We consider various combinations of $\delta$ and $\lambda$ ; $\delta =$( 0, 0.25, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4) and $\lambda =$ (0, 0.5, 1, 1.5, 2, 3). We report the results for $p = 2$ for warmup series (10, 20, 30, 40). The results can be summarized as follow:

– The effect of warmup series:

They seem to matter most when the shift in mean is of small magnitude; the results suggest that longer warmup series yield faster responses to small shifts. For big shifts in mean (i.e. in the magnitude of $\delta \geq 3$,) the system has a fast response to shift regardless of the warmup series sizes

($4 \leq$ OOC.ARL $<10$ for all $\delta \geq 3$ at all combinations of $\lambda$ and warmup series lengths).

– The effect of mean shift:

The bigger the shift in mean vector, the faster the response. There was a relative robustness to location parameter detection under basic performance study.

– The effect of shift in covariance matrix:

Under basic performance study, there was an overall faster response in the presence of a sudden change in the sum of squares; the breaking points, in dimension 2, comes when $\lambda > 1.5$.

– Small shifts in mean vector coupled with small shift in variance do not have a fast detection. Not a surprise; they are the closest to in-control state. Figure 1 below give a clearer picture on the performance.

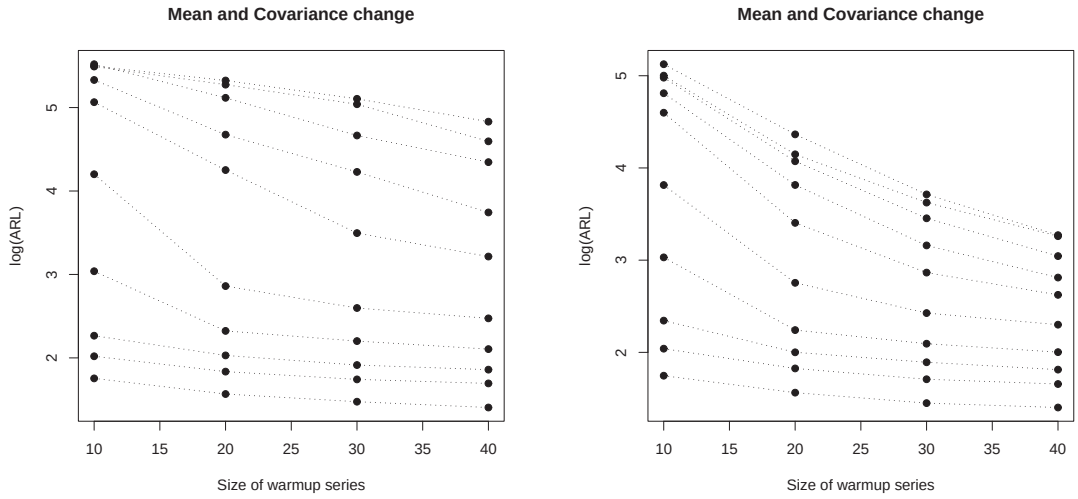


Figure 1: Basic performance study: $\delta =$( 0, 0.25, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4) and $\lambda =$ (1.5(left) , 3(right)) warmup series are 10, 20, 30, 40. Each dotted line represents a specific mean shift; starting with 0 at the top and 4 at the bottom. The vertical axes are the log(ARL).

When examining the spread of the OOC runs though, performances were more erratic in the presence of unreasonably small size data gathering. The spread of the runs stabilizes around warmup series of size 20 when $p = 2$.

## Performance under More General Covariance Changes

The covariance matrix contains $p(p+1)/2$ elements, so a full exploration of all possible shifts is a daunting task. It is though not quite as daunting as it looks initially. The methods of this paper are affine equivariant – they depend on the covariance matrices $\boldsymbol{\Sigma}, \boldsymbol{\Sigma}_1$ only through the eigenvalues of $\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}^{-1}$, or, with the standardization $\boldsymbol{\Sigma} = \boldsymbol{I}$, those of the shifted covariance matrix $\boldsymbol{\Sigma}_1$. Multiplying $\sigma_{1,1}$ by $1 + \lambda$ replaces the eigenvalues $1, \ldots, 1$ with $1 + \lambda, 1, \ldots, 1$, and in fact captures the effect of any rank-1 update to $\boldsymbol{\Sigma}$. This rank-1 update is equivalent to the basic performance discussed previously, in which no correlation has been introduced between quality characteristics. A more general form of all rank-1 updates can be written as changing the covariance matrix from identity to a matrix with first entry $\sigma_{1,1}^2 \neq 1$; in which case proportional reduction in variance can be explored by constraining $0 < \sigma_{1,1}^2 < 1$; which in turn can be mapped on $\lambda \in (-1, 0)$. Note though, that this scenario is not necessarily restricted to a single component update; it can also be viewed as all changes that correspond to a rank-1 update from $\boldsymbol{I}$.

Introducing a non-zero correlation $\varrho$ between two elements of $\boldsymbol{X}$ changes the eigenvalues to $1 + \varrho, 1 - \varrho, 1, \ldots, 1$. This is a rank-2 update, though of a rather special form.
The reason for exploring more general covariances is that, in multivariate settings, it is customary to study changes from $\boldsymbol{I}$ to a covariance matrix with some particular structure with regard to their eigenvalues. This full exploration will require more space than the current manuscript and would probably detract from the objective of the paper. We consider three different settings of rank–2 updates, where the performance of the chart following a loss of control is inextricably linked to changes in eigenvalues of the covariance matrix following the change point. We carried out a study in dimension $p = 2$; although the idea

is generalizable to higher dimensions. We consider three different settings called settings **1.**, **2.**, **3.**, and coupled them with a constant change in mean. The various covariance structures are:
$\begin{pmatrix} 1-\varrho & 0 \\ 0 & 1-\varrho \end{pmatrix}$, $\begin{pmatrix} 1 & \varrho \\ \varrho & 1 \end{pmatrix}$, and $\begin{pmatrix} 1+\varrho^2 & \varrho \\ \varrho & 1+\varrho^2 \end{pmatrix}$. We ran these three scenarios following 10, 20, 30, and 40 learning series for a system tuned to ARL = 500. We chose $\varrho = \{.1, .2, .4, .6, .8\}$; and various contaminations of the mean $\delta = \{0, .25, .5, 1, 2, 3, 4\}$. For each setting, we ran 10,000 replications of series of length 3000 each and looked for signal. The results are included on Tables 2, 3, 4, and on Figures 2 & 3.

**1.** $\begin{pmatrix} 1-\varrho & 0 \\ 0 & 1-\varrho \end{pmatrix}$ explores a system switching from $\boldsymbol{I}_2$ to a system with equal reduction (or equal % reduction) in variance. In dimension 2 this yields a change in eigenvalues from 1, 1 to 1- $\varrho$, 1- $\varrho$. In dimension $p$, this change can define a rank-2 update from $\boldsymbol{I}$ to $\boldsymbol{\Sigma}_1$, when two of the entries have equally affected the overall variability; consequently changing the eigenvalues to $1-\varrho, 1-\varrho, 1, \ldots, 1$ . The results of this setting (Figures 2, 3, & Table 2) show that our system responds faster to moderate-to-large decreases in the variance component. When this decrease is tiny ($\varrho \leq 0.2$; nearly identical curves across the range of $\delta$), the system can go undetected for a long time (out of control(OOC) ARL $\sim$ 430 when $\delta = 0$); unless a large shift in mean comes to its rescue (OOC.ARL $\sim$ 4 when $\delta = 4$). As the shift in mean increases and the decrease in variance becomes apparent, the change point method is more likely to flag changes at a faster pace. For large shifts in mean vector ($\delta \geq 2$), the method has a fast detection and seems to behave so with the startup series having little effect (OOC.ARL decreasing from 11.1 ($n_0 = 10$, $\delta = 0$ , $\varrho = .1$ ) to 3.9 ($n_0 = 40$, $\delta = 4$ , $\varrho = .8$)).

(On a side note, this scenario and a rank-1 update with decrease in variance in view can be obtained from each other and interchangeably. This is not difficult to see; in one case the determinant is the square power of the other. So, it would be redundant to add an additional simulation study on rank-1 variance decrease to the paper.)

**2.** $\begin{pmatrix} 1 & \varrho \\ \varrho & 1 \end{pmatrix}$ explores a system switching from $\boldsymbol{I}_2$ to a correlated system. In dimension 2, this yields a change in eigenvalues from 1, 1 to 1 - $\varrho$ , 1 + $\varrho$. In dimension $p$, this change can define a rank-2 update from $\boldsymbol{I}$ to $\boldsymbol{\Sigma}_1$ when one update increases the overall variability while the other decreases it. This changes the eigenvalues to 1 - $\varrho$ , 1 + $\varrho, 1, \ldots, 1$. Figures 2, 3, & Table 3 show the performance of the change point model under this scenario in dimension 2. When a correlation creeps into the system, its magnitude plays a key role in it being detected. The change point method is fast to detect a process changing from uncorrelated to highly correlated. We experience a slow response when $\varrho \leq 0.4$ and $\delta \leq .5$. The OOC.ARL ranges from around 400 ($\varrho = .1$, $\delta = 0$; $n_0 = 10$) to 112 ($\varrho = .4$, $\delta = .5$; $n_0 = 40$). When a change in correlation is coupled with a relatively large shift in the mean ($\delta \geq 2$), the size of the learning observations appears to have little to no importance. The OOC.ARL ranges from 11.8 ( $\varrho = .1$, $\delta = 2$; $n_0 = 10$) to 4.0 ( $\varrho = .8$, $\delta = 4$; $n_0 = 40$).

**3.** $\begin{pmatrix} 1+\varrho^2 & \varrho \\ \varrho & 1+\varrho^2 \end{pmatrix}$ explores a more general scenario where both variance and correlation are subject to changes. In dimension 2, this yields a change in eigenvalues from 1, 1 to $1 + \varrho^2 + \varrho$, $1 + \varrho^2 - \varrho$. The system switched from the uncorrelated $\boldsymbol{I}_2$ to a correlated system with slight augmentation in variance. In dimension $p$, this generalizes to a rank-2 update in which eigenvalues change to $1 + \varrho^2 + \varrho$, $1 + \varrho^2 - \varrho$, 1, ... 1– representing a target intermediate between variance and correlation shifts. Figures 2, 3 & Table 4 show the performance of the change point model under this scenario in dimension 2. This setting has the slowest responses across the range of $\varrho$ when $\delta$ is tiny, and the fastest responses when $\delta \geq 2$, regardless of the startup observations. The figures show little separation between between the various $\varrho$ at a small shift in mean, a sharp decrease in OOC.ARL ($\delta \geq 2$) with no separation across $\varrho$. For large $\delta$, the OOC.ARL ranges from 12 ( $\varrho = .1$, $\delta = 2$; $n_0 = 10$) to 4.1 ( $\varrho = .8$, $\delta = 4$; $n_0 = 40$).

Overall, all three settings are nearly equivalent when $\varrho \to 0$ and $\delta \to 0$, and are expected to draw close to their nominal ARL = 500. The farther $\varrho$ and $\delta$ get from 0, the faster the charts respond to shifts. In setting **2.**, when a correlation creeps into the process and $n_0$

is small, the change can go undetected for some time; unless the correlation is significantly high or a shift in the mean comes to its rescue. The detection gets a little faster when $n_0$ is large. Finally, when the system goes out of control as specified in **3.**, we observe the slowest responses compared to **2.**; especially when drifts in variance and correlation are of very small magnitude and the mean shift negligible. By and large, the system is more sensitive to moderate to large increases in mean ($1 < \delta \leq 4$), couple with a significant changes in correlation ($\varrho \geq .4$).

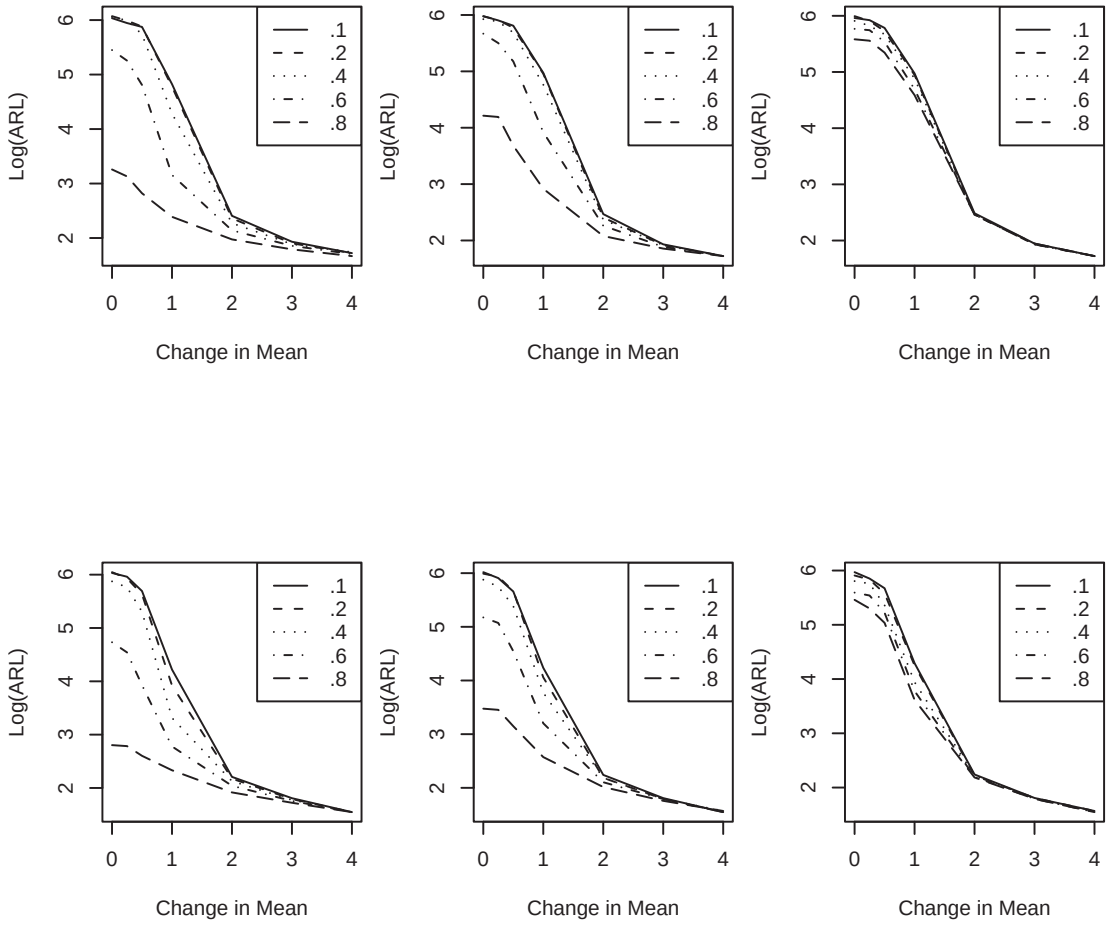


Figure 2: Performance under more general covariance changes. From left to right: settings **(1. ; 2. ; 3.)**, First row : **Startup 10**, Second row : **Startup 20**; $\varrho = (.1, .2, .4, .6, .8)$ on legend.

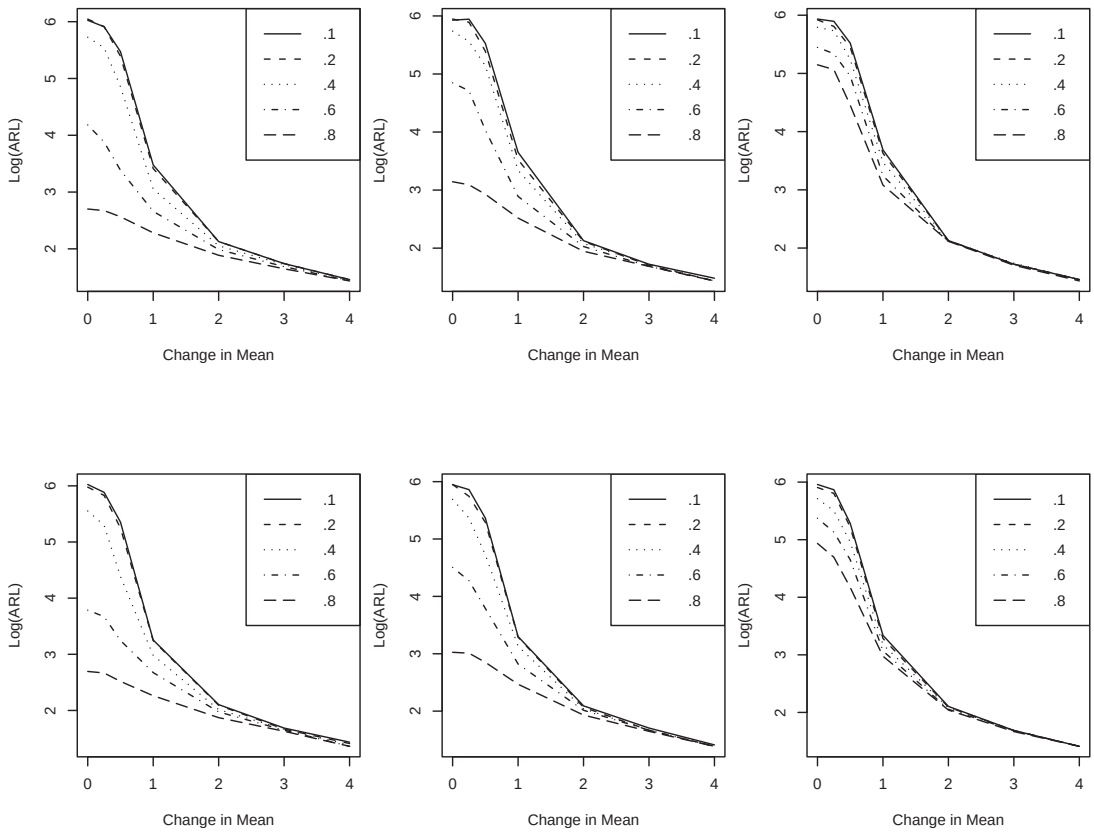


Figure 3: Performance under more general covariance changes. From left to right: settings **(1. ; 2. ; 3.)**, First row : **Startup 30**, Second row : **Startup 40**; $\varrho$ = (.1, .2, .4, .6, .8) on legend.

# Computational Issues

In an era of advanced computational technology, one may be tempted to implement our methodology in a crude or *'brute-force'* way. However, direct implementation of the formulas involved in the generalized likelihood ratio criterion, and their sequential computation can lead to much more severe computational load than is necessary. The criterion depends on the computation of many determinants $\mid \boldsymbol{S}_{i,j} \mid$ of the covariance matrices of subsequences of the data. These determinants can be found using a numerically stable reasonably fast method based on Cholesky factorizations.

Append a 1 to each data vector $\boldsymbol{X}_h$, writing

$$\boldsymbol{Z}_h = (1, \boldsymbol{X}'_h)' \tag{8}$$

Define

$$\boldsymbol{V}_{i,j} = \sum_{h=i+1}^{j} \boldsymbol{Z}_h \boldsymbol{Z}'_h \tag{9}$$

and form the lower triangular Cholesky factorization

$$\boldsymbol{V}_{i,j} = \boldsymbol{R}_{i,j} \boldsymbol{R}'_{i,j} \quad ; \tag{10}$$

where the Cholesky factor matrix $\boldsymbol{R}_{i,j}$ is of dimension $p+1$. Standard results then show (see for example Chambers 1971) that

$$|\boldsymbol{V}_{i,j}| = \prod_{h=2}^{p+1} r^2_{h,h,i,j} \quad ; \tag{11}$$

where $r_{h,h,i,j}$ is the $h,h$ element of $\boldsymbol{R}_{i,j}$.

The attraction of the Cholesky factorization is that adding a new data vector $\boldsymbol{X}_{j+1}$, the corresponding $\boldsymbol{R}_{i,j+1}$ can be computed from $\boldsymbol{R}_{i,j}$ and $\boldsymbol{Z}_{j+1}$ with a fast, stable update (see Chambers 1971). Starting off for each $i = 1, \ldots$ with $\boldsymbol{R}_{i,i} = 0$ and then using updates to calculate the successive $\boldsymbol{R}_{i,j}$, all the log determinants needed for the analysis can be calculated. These can be computed "on the fly" as needed, which involves minimal storage requirement but at the cost of much repetitive computation. A more efficient approach though is to keep a record of all determinants $Q(i,j) = \log |\boldsymbol{V}_{i,j}|$, (it is better to work with the logs of the determinants than the determinants, as the latter are at risk of overflow and underflow in high dimensions.) If all $Q(i,j)$ for $0 \leq i < j \leq n$ have been computed, then when case $n+1$ accrues, initialize a temporary working Cholesky factor $\boldsymbol{R}_{n+1,n+1} = 0$, and starting from $i = n+1$ and going down to $i = 0$ use the updates to successively compute each $\boldsymbol{R}_{i-1,n+1}$, and from this, $Q(i, n+1)$.

There are notionally $n(n-1)/2$ elements of $Q$, though those for which $j < i+p$ are necessarily $-\infty$ and can be ignored.

# Application: Detection of signs of 'Isolated Systolic Hypertension'(ISH)

This problem was the motivation that has partly driven the research work. It also reemphasizes the call in the U.S. and around the globe to apply advanced quality methodologies to health science, in order to advance symptoms recognition and to improve health care systems; (Quality and Productivity Research Conference(QPRC), 2005, Minneapolis, MN). Authors such as Woodall (2006), Winkel & Zhang(2007), have responded favorably to the call to apply quality control methods in health and in medicine. In fact, Woodall(2006) has conducted works that pertain to applying industrial quality control methods to health science. The current application of our methodology falls within the same line of thoughts–with hope that many authors will continue to respond to this urgent call.

In chronobiology and medical ambulatory monitoring, investigators are interested in the behavior of subjects measured over time, on some specific biological characteristics, in order to assess risks. As an example, monitoring systolic and diastolic blood pressures can tell whether a person is at risk for stroke, death, or other medical complications and disabilities associated with heart diseases. As individuals age their systolic blood pressure has a tendency to rise, while the diastolic pressure falls. This combination can result in one of the dominant disease processes in people aged 50 years or older called *'Isolated Systolic Hypertension (ISH)'*. This comes about when the systolic blood pressure is high (usually between 120–140 mm Hg), while the diastolic blood pressure remains below 90 mm Hg. This is in contrast to traditional definitions of hypertension in which high blood pressure is defined as both the systolic and diastolic components crossing their hypothetically specified thresholds. With $ISH$ the pressure in the arteries may become very low when the heart relaxes; resulting in lower diastolic blood pressure. The latest medical researches have suggested that both systolic and diastolic measures be taken into account before classifying patients as high blood pressure individuals, or as high/low risk for heart disease. $ISH$ is, in and of itself, considered a risk for heart disease, and warrants medical attention; though it may be further

complicated by other co-morbidities. It has also been recognized that in normal subjects, systolic and diastolic blood pressures are highly correlated. Thus, monitoring these biological characteristics separately without accounting for the correlation between them would be discordant with process control realities. In the following study, we have data set on a subject over 50 years of age, oscillometrically measured over time. The goal is to identify, as soon as possible, signs of $ISH$ using single observations/vectors obtained sequentially and periodically. The setting can be viewed as a bivariate sequential/dynamic control problem with correlated characteristics. Some may argue that in two dimension settings such as this one, only five parameters are in play (2 means, 2 variances and one covariance); and that life is easier using univariate methods on these parameters. We do not choose this *'comfort zone'* alternative; since we are more concerned with a joint charting methodology as opposed to separate univariate charts. We apply the multivariate change point control methodology to these data to pinpoint the major time of change; bearing in mind that change can occur in both the mean and/or the covariance matrix.

## Result of change point method

Figure 4 shows the results of applying the change point methodology to our data using two different settings. The first setting accumulates 10 observations as initial sample then starts testing from observation 16, while the second uses 20 initial samples and starts testing at observation 26. On the left-hand panel, the 'nearly horizontal' lines on the figures are the control limits $h_{n,0.002,2}$, corresponding to an ARL of 500 (Table 5), and the curves are the test statistics. The right-hand panel of each figure shows the maximum likelihood estimate of the epoch of change ($\hat{\tau}$). The message conveyed by the figures is essentially identical, except the short-warmup sequence's flirtation with the control limit around observation 30. In all, the test statistic punches through the control limit after adding the $55^{th}$ observation, remains above the limit for the rest of the sequence. From observation 55 on, the estimate of the change point remains fixed and consistent at observation 48.

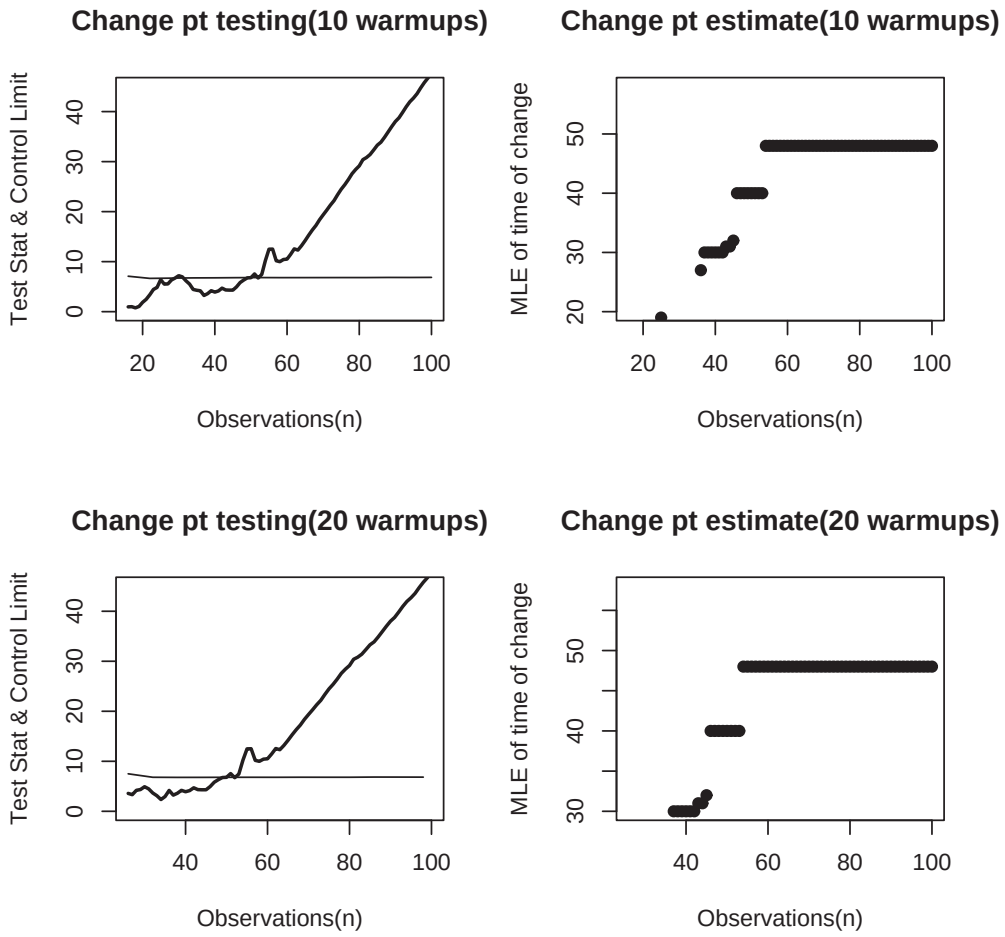


Figure 4: Test statistics, Control Limits & Change point Estimates

## Diagnosis after Signal

In multivariate control setting, it is important to carry a diagnosis after signal to see what characteristics of the process have gone out-of-control. Unlike the univariate setting, this investigation may be quite complex. In our case, since the MLE estimate $\hat{\tau}$ suggested a major split at time point 48, a closer look at the pre and post-shift data revealed the following summaries:

$$\boldsymbol{\mu}_{1\ldots\hat{\tau}} = \begin{pmatrix} 125.16 \\ 78.77 \end{pmatrix} \qquad \Sigma_{1\ldots\hat{\tau}} = \begin{pmatrix} 8.54 & 3.32 \\ 3.32 & 8.46 \end{pmatrix} \qquad \texttt{Corr}_{1\ldots\hat{\tau}} = \begin{pmatrix} 1.00 & 0.39 \\ 0.39 & 1.00 \end{pmatrix};$$

$$\boldsymbol{\mu}_{\hat{\tau}+1\ldots55} = \begin{pmatrix} 129.03 \\ 76.94 \end{pmatrix} \quad \Sigma_{\hat{\tau}+1\ldots55} = \begin{pmatrix} 7.23 & 4.83 \\ 4.83 & 3.26 \end{pmatrix} \quad \texttt{Corr}_{\hat{\tau}+1\ldots55} = \begin{pmatrix} 1.00 & 0.99 \\ 0.99 & 1.00 \end{pmatrix}.$$

The first component of each mean vector represents the mean of systolic measurements. As a matter of interest, we may also compare the post-shift parameter estimates obtained at the time of the signal, observation 55, with the estimates given at the end of the data sequence; the two sets of estimates are encouragingly similar.

$$\boldsymbol{\mu}_{\hat{\tau}+1\ldots n} = \begin{pmatrix} 128.25 \\ 77.75 \end{pmatrix} \quad \Sigma_{\hat{\tau}+1\ldots n} = \begin{pmatrix} 9.40 & 5.32 \\ 5.32 & 4.17 \end{pmatrix} \quad \texttt{Corr}_{\hat{\tau}+1\ldots n} = \begin{pmatrix} 1.00 & 0.84 \\ 0.84 & 1.00 \end{pmatrix}.$$

These summaries suggest that the signal may have resulted from a shift in the systolic blood pressure and a distortion in the correlation structure from weak to highly correlated variables, coupled with a mean and/or variance reduction in the diastolic blood pressure measurements. In order to justify this claim, we use the step-down analysis for change as outlined in Sullivan et al.(2007), although there are other proposals for the diagnosis following signals such as Hawkins and Maboudou-Tchao (2008).

**Step-down Method**

The reasoning behind the step-down method consists of using the difference in estimated parameter vectors as diagnosis tool to infer which elements of the difference are not zero. If the parameter vector is labeled $\theta = (\mu_{sys}, \mu_{dia}, \sigma_{sys}, \rho_{(sys,dia)}, \sigma_{dia})$, the idea is to partition the difference in parameter vectors $\Delta = (\theta_{pre} - \theta_{pos})$ into a test subset and its complement (i.e. $[\Delta_{in} = (\theta_{pre} - \theta_{pos})_{in} \mid \Delta_{out} = (\theta_{pre} - \theta_{pos})_{out}]$). The test subset $(in)$ consists of the parameters (or combination of parameters) we wish to test for homogeneity between the $pre-$ and $post-$ shift series. If we wish to test for homogeneity of the systolic component regardless of the other four parameters of the space, the design vector (1, 0, 0, 0, 0) instructs that the systolic parameter $\mu_{sys}$ is in the test subset $\Delta_{in}$ while the other four are not; thus, they are in $\Delta_{out}$. The design matrix of Table 6, explains all there is to be known about partitioning $\Delta$ into $\Delta_{in} \bigsqcup \Delta_{out}$. The design vector (1, 0, 1, 0, 0) tests for the mean and the variance of the systolic measures; the other components of the bivariate measures, i.e., the

diastolic parameters and the correlation between the two characteristics will be in the $\Delta_{out}$ set. The test of homogeneity, for a parameter in the test subset, is testing the null of no difference between pre-change and post-change measurements. By varying the test subset, one can cover all potential combination of parameter changes and find the best subset of parameters that explains the change we have flagged earlier. The technicality partitions the estimated covariance matrix of the difference into

$$\hat{\mathbf{\Sigma}}_{\hat{\Delta}} = \begin{pmatrix} \hat{\Sigma}_{in} & \hat{\Sigma}_{1,2} \\ \hat{\Sigma}'_{1,2} & \hat{\Sigma}_{out} \end{pmatrix}, \tag{12}$$

and decomposes the overall quadratic $\chi^2 = \hat{\Delta}'\hat{\Sigma}^{-1}_{\hat{\Delta}}\hat{\Delta}$ into $\chi^2 = \chi^2_{in} + \chi^2_{out|in}$, where $\chi^2_{in} = \hat{\Delta}'_{in}\hat{\Sigma}^{-1}_{in}\hat{\Delta}_{in}$; $\chi^2_{out|in} = \hat{\Delta}'_{out|in}\hat{\Sigma}^{-1}_{out|in}\hat{\Delta}_{out|in}$; $\hat{\Delta}_{out|in} = \hat{\Delta}_{out} - \hat{\Sigma}'_{1,2}\hat{\Sigma}^{-1}_{in}\hat{\Delta}_{in}$; and $\hat{\Sigma}_{out|in} = \hat{\Sigma}_{out} - \hat{\Sigma}'_{1,2}\hat{\Sigma}^{-1}_{in}\hat{\Sigma}_{1,2}$.

$\chi^2_{in}$ is asymptotically central chi-squared distributed, and $\chi^2_{out|in}$ is asymptotically non-central chi-squared distributed. The asymptotic covariance estimate for the blood pressure problem under the parameterization $(\mu_{syst}, \mu_{dias}, \sigma_{syst}, \rho_{(syst,dias)}, \sigma_{dias})$ is

$$\hat{\mathbf{\Sigma}}_{\hat{\Delta}} = (1 + \tfrac{48}{8}) \begin{pmatrix} 0.17791670 & 0.06906174 & 0.0 & 0.0 & 0.0 \\ 0.06906174 & 0.17625000 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.08895833 & 0.01006623 & 0.01346704 \\ 0.0 & 0.0 & 0.01006623 & 0.01497780 & 0.01001897 \\ 0.0 & 0.0 & 0.01346704 & 0.01001897 & 0.08812500 \end{pmatrix}.$$

The pre and post-shift estimated parameters are

$$\hat{\theta}_{\texttt{pre}} = (125.16, 78.77, 2.922328, 0.39, 2.908608)'$$

$$\hat{\theta}_{\texttt{pos}} = (129.03, 76.94, 2.688866, 0.99, 1.805547)'.$$

Table 6. shows the step-down method for all possible parameter tests. In the table, $n_p$ refers to the number of parameters tested. The ones stand for parameters included in the test

set while the zeroes are the parameters excluded from the set. The final decision as to which parameters have changed individually or jointly depends on a pre-specified decision making threshold. The choice of the decision threshold ($p^*$) has sensitivity implications. It is chosen to neither identify too many changes nor too few. In other words, it is chosen to have a good sensitivity to identify the correct number of parameter change. Any observed significance that falls below $p^*$ gives a convincing evidence that the combination of parameters tested has changed. If in a given $n_p$, the maximum observed significance across all combinations is below $p^*$, we would have evidence that the ideal number of parameter change is $n_p$. Following a reasoning similar to Sullivan et al.(2008), we set $p^* = 0.20$ as decision making threshold.

When $n_p = 1$, the parameter most consistent with the assumption of 'no change' is $\sigma_{sys}$. There is a convincing evidence that the rest of the parameters might have changed.
When $n_p = 3$ or 4, all subsets are statistically inconsistent with the assumption of no change. Identifying the smallest $n_p$ for which all p-values are less than $p^*$ is key to identifying the ideal number of parameters violating the assumption of no change. We identify the smallest value of $n_p$ to be 3; although one can reasonably have the same argument about $n_p = 4$.

The conclusions from the change point detection algorithm and the step-down method have led us to partitioning the readings as displayed in Figure 5. From the figure, we see a weak association between SBP and DBP up to $\hat{\tau}$; after which the association becomes significantly strong. In addition, the time series plot suggested some irregular behavior between these measurements up till around time point 48 (one characteristic increasing in variance while the other decreases). By and large, our conclusion based on the step down algorithm is not far from this visual inspection that the change is a by-product of a highly significant increase in systolic blood pressure after time point 48 (which is a 1.32 standard deviation increase), an increase in the correlation between the two blood pressures (correlation increases by a factor of more than 2.5), a .63 standard deviation decrease in diastolic blood pressure–although one can also argue about a deflation in variance of the diastolic blood pressure. Statistically, we have identified the epoch of change and the possible direction of the change. In medical

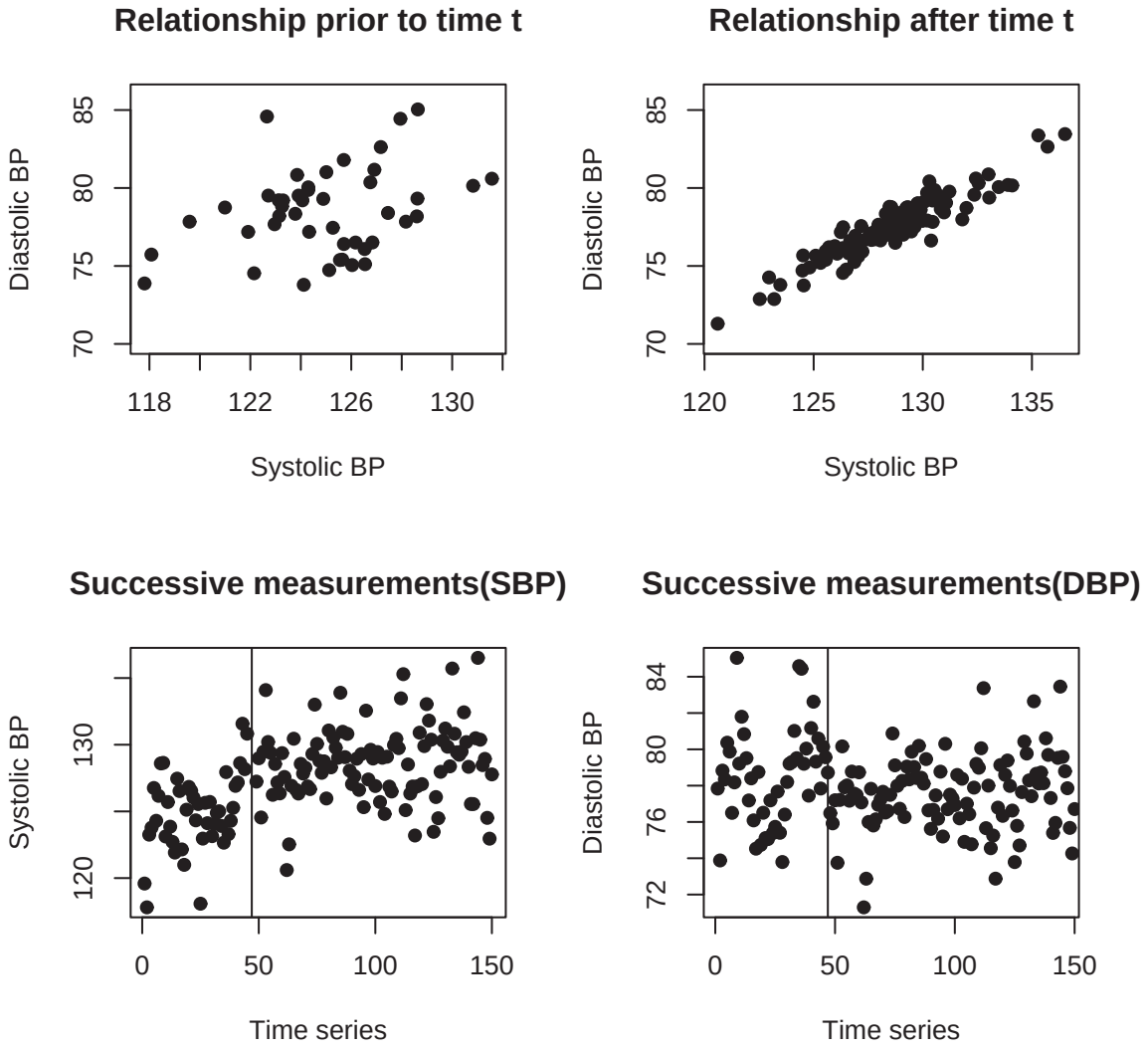


Figure 5: Blood pressure problems: Joint relationship and marginal time assessments.

fields, statistical significance does not necessarily translate into clinical importance. In our setting though, the statistical diagnoses have a two-fold implication to medical doctors and ambulatory monitors. The first implication is to identify whether a regimen is having a stabilizing factor on the main biological variables, while the second identifies when an $ISH$ patient should be placed on a blood pressure stabilizing regimen. These indeed have direct repercussions on assessing risk for heart diseases within this population.

# Comparative Study

In advancing a new methodology, one wishes to compare it to an existing standard. The existing standard in our case is lacking due to the fact that standard control techniques rely on known parameters. Reynolds & Cho (2006) comes to mind as a possible comparator; their method, however, cannot accommodate variance decreases or correlation changes, limiting its relevance. The closest benchmark for comparison is to use a heuristic approach by designing univariate self-starting cusums of regression recursive residuals. Hawkins & Maboudou-Tchao (2007) points to the transformation to self-starting regression-adjusted variables as creating a stream of quantities that one can use for any charting purposes. The method transforms an original unknown parameter vector $\boldsymbol{X}$ into recursive residuals $\boldsymbol{R}$, then transforms $\boldsymbol{R}$ into known parameter multivariate standard normal vector $\boldsymbol{U}$ so that changes in $\boldsymbol{X}$ are reflected upon $\boldsymbol{U}$. Thus, our approach is to set up $p$ univariate streams on these recursive residuals which, in control, are iid $N(0,1)$, and maintain a pair of cusums for location (one upward and one downward) and a pair for scale on each stream, for a total of $4p$ cusums that are run until one of them crosses its decision interval, so that the run length(RL) is the $\min(\text{OOC.RL}_i)_{i=1}^{4p}$. We compare this first-signal-based cusum to the performance of the change point (CP) methodology. As you can see, these comparisons can become burdensome in high dimension. We demonstrate this for $p = 3$ only. In order to obtain a joint in-control ARL of 500, we must tune each individual self-starting cusum to a nominal ARL of 6000. We consider change in the covariance structure on the form $\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{1,1}^2 & 0 & 0 \\ 0 & \sigma_{2,2}^2 & 0 \\ 0 & 0 & \sigma_{3,3}^2 \end{pmatrix}$. The reference value and the decision intervals for the cusums depend on the shift we anticipate to detect. In this simplistic demonstration, we operate under the assumption that each of the parameters $(\mu_1, \mu_2, \mu_3, \sigma_{1,1}, \sigma_{2,2}, \sigma_{3,3})$ can increase, decrease, or remain stable. This clearly presents a combinatorial explosion of comparisons if we were to study them individually. We proceeded randomly; since, in the real world, we would not know exactly where a shift may be coming from. In addition, what we need to prioritize is the fast response to shifts rather

than response to specific shifts. Each random selection of these combinations is run through the change point method and through the 12 cusums. At least one shift was introduced after 25 in-control readings. The table below (Table 7.) gives us a design scheme for the specific shifts to be detected by the cusums. The OOC.ARL of the table are the out of control ARL of the cusum design schemes. We measure performance by the number of series signaling within a caliper defined by 100 observations after the in-control readings. So, any series that did not signal after 100 contaminations is considered degenerate and is censored. This is a heuristic approach in which we are comparing truncated out of control ARL. The OOC.ARL for the comparison are based on 10,000 replications. The results are a little surprising. The OOC.ARL(Cusum) = 20.09, and the OOC.ARL(CP)= 9.39; giving a clear advantage to the CP method, for fast detection, over multiple cusums. Note too, that out of 10,000 randomly generated series, 2.7% were degenerate for the change point whereas 2.5% were degenerate for the cusums. One surprise from our simulation though, is that only a few upward scale cusums signaled (9.68%); and that location cusums dominated the signals (61%). Even though this looks at first blush a little unrealistic, it confirms the fact that one needs to be very careful when using univariate charts to monitor multivariate characteristics. The reality is, when using startup cusums, shifts in scale parameter can easily reflect upon the location parameter cusum and the running mean can easily get caught up with newly shifted parameters. Figure 6 shows the side-by-side histograms for these comparisons.

**Table 7.**

| | Shift | $h$ | $k$ | ARL | OOC.ARL |
|---|---|---|---|---|---|
| `Location` | $1\sigma$ Increase | 6.852 | 0.50 | 6004.80 | 14.1 |
| | $1\sigma$ Decrease | 6.852 | 0.50 | 6004.80 | 14.1 |
| `Scale` | $0.5\sigma$ Increase | 20.868 | 1.460 | 6000.40 | 27.3 |
| | $0.5\sigma$ Decrease | 4.683 | 0.462 | 6002.10 | 22.0 |

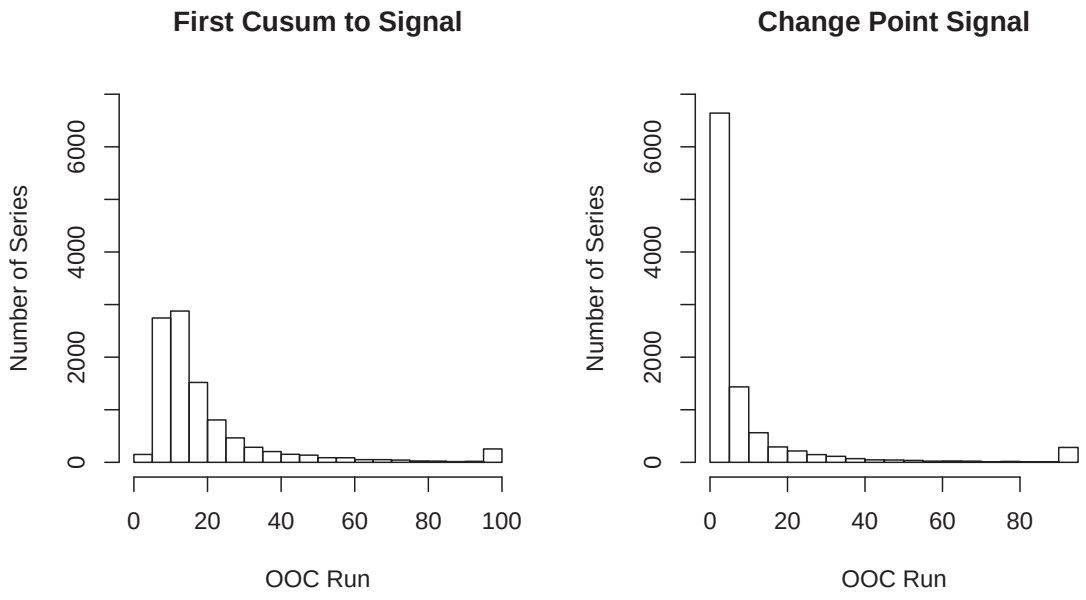


Figure 6: Comparison of the change point method to univariate startup cusums. The comparison was carried in dimension 3 and 4×3 cusums were used. Means and variances are allowed to shift according to the design table 7, or remain stable.

# Discussion & Conclusion

Traditional charting systems are calibrated under the assumption that the in-control process parameters are known exactly. To the extent that this assumption became violated, the in-control run behavior of the resulting chart will differ from what the user expected. In addition, under traditional control settings, it is necessary to gather large Phase I data sets in order to set the process parameters. Cusum and MEWMA have a further challenge in that their best performance requires 'tuning' to the size of the shift. Our unknown-parameter change point formulation removes the need for a large Phase I exercise. While it is probably

a good idea to gather some Phase-I-type familiarization data before starting formal process monitoring, it is not necessary for this initial set to be large, and in the extreme case of short-run processes where prior information indicated a likely normal distribution for the process readings, it could be skipped entirely, and monitoring started with the $2(p+1)^{th}$ process reading. Apart from the potential for greatly reducing the amount of effort put into gathering Phase I data, we see a large benefit of the change point formulation in that it removes much of the distinction between Phase I and Phase II, allowing process learning to continue beyond the usual Phase I boundary in an automated, seamless way.

Unknown-parameter methods such as this one can be viewed either as a complete package, or as a stop-gap en route to the known-parameter setting. In line with this, some users might want to use the change point formulation until enough in-control data had been gathered to effectively remove all estimation error from the parameters, at which point they would go over to their favorite known-parameter charting method – for example a cusum or MEWMA. Other users would continue to use the change point method indefinitely, an approach that has the virtue of not involving some sharp break where we switch from one monitoring approach to another. We are inclined toward the latter approach.

Finally, there is the matter of the Phase I / Phase II dichotomy. We believe the change point approach allows one to blur the normally-sharp line between these two, allowing monitoring during the early part of data gathering, along with continuous improvement in parameter estimates during the ongoing data gathering. The main task for the Phase-I-like portion of the exercise is not to estimate the parameters, but to get a feeling for the distribution of the process vectors.

# Appendix

Assuming a split point at time $k$, the generalized likelihood ratio test statistics for a series of $n$ observations from a $p$-variate normal distribution is:

$$\Lambda_{k,n} = \frac{\mid \hat{\Sigma}_{0,k} \mid^{\frac{k-1}{2}} \times \mid \hat{\Sigma}_{k,n} \mid^{\frac{n-k-1}{2}}}{\mid \hat{\Sigma}_{0,n} \mid^{\frac{n-1}{2}}}$$

$$\begin{aligned} -2\log(\Lambda_{k,n}) = & \ (n-1)\log \mid \hat{\Sigma}_{0,n} \mid -(k-1)\log \mid \hat{\Sigma}_{0,k} \mid -(n-k-1)\log \mid \hat{\Sigma}_{k,n} \mid \\ = & \ C + (n-1)\log \mid \boldsymbol{S}_{0,n} \mid -(k-1)\log \mid \boldsymbol{S}_{0,k} \mid -(n-k-1)\log \mid \boldsymbol{S}_{k,n} \mid \end{aligned} \tag{13}$$

where $C = p\,[\log(n-1) - \log(k-1) - \log(n-k-1)]$.

The distribution of the generalized variance of a sample $X_1, \ldots, X_n$ from $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is the same as the distribution of $\mid \boldsymbol{\Sigma} \mid /(n-1)^p$ times the product of $p$ independent factors, the distribution of the $j-th$ factor being the $\chi^2$- distribution with $n-j$ degrees of freedom; Theorem 7.5.3 (Andersen 1984).

It follows that $\mid \boldsymbol{S}_{0,n} \mid$ is distributed as

$$\begin{aligned} \mid \boldsymbol{S}_{0,n} \mid \ & \sim \ \mid \Sigma_{0,n} \mid \times \chi_{n-1} \times \chi_{n-2} \times \ldots \times \chi_{n-p} \\ & \sim \chi_{n-1} \times \chi_{n-2} \times \ldots \times \chi_{n-p}; \end{aligned}$$

(assuming w.l.o.g that $\boldsymbol{\Sigma} = \boldsymbol{I}_p$).

$$\begin{aligned} \log \mid \boldsymbol{S}_{0,n} \mid \ & \sim \log(\chi_{n-1}) + \log(\chi_{n-2}) + \ldots + \log(\chi_{n-p}) \\ & \sim \log(W_1) + \log(W_2) + \ldots + \log(W_p) \end{aligned}$$

and

$$E(\log \mid \boldsymbol{S}_{0,n} \mid) \ = \ E(\log(W_1)) + E(\log(W_2)) + \ldots + E(\log(W_p));$$

where $W_j \ \sim \chi^2_{n-j}$.

If $M_{(\log W_j)}(.)$ is the moment generating function of the $j-th$ term,

$$M_{(\log W_j)}(t) = E(e^{t \log W_j}) = E(W_j^t)$$

$$\frac{1}{\Gamma(\frac{n-j}{2}) 2^{(\frac{n-j}{2})}} \int w^t w^{(\frac{n-j}{2})-1} e^{\frac{-w}{2}} dw \ = \ \frac{\Gamma(t + \frac{n-j}{2})}{\Gamma(\frac{n-j}{2})} \ 2^t$$

$$E(\log(W_j)) \ = \ \lim_{t \to 0} \frac{d}{dt} M_{(\log W_j)}(t) \ = \ \lim_{t \to 0} \frac{d}{dt} [\frac{\Gamma(t + \frac{n-j}{2})}{\Gamma(\frac{n-j}{2})} \ 2^t] \ = \ \log 2 + \psi(\frac{n-j}{2})$$

thus,

$$E(\log \mid \boldsymbol{S}_{0,n} \mid) \rightarrow p \ \log 2 \ + \ \sum_{j=1}^{p} \ \psi(\frac{n-j}{2}). \tag{14}$$

For an assumed split point at $k$, summing $\boldsymbol{S}_{0,k}$ , $\boldsymbol{S}_{0,n}$ and $\boldsymbol{S}_{k,n}$ of (13) gives the analytical result.

**Table 1.**

Control Limits $h_{n,\alpha=0.002,p}$

| | $p$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 2 | 3 | 4 | 5 | 10 | 15 | 20 | 25 |
| $n$ | | | | | | | | |
| 16 | 7.089 | | | | | | | |
| 17 | 6.785 | | | | | | | |
| 18 | 6.686 | 4.474 | | | | | | |
| 19 | 6.647 | 4.493 | | | | | | |
| 20 | 6.648 | 4.506 | 3.549 | | | | | |
| 21 | 6.648 | 4.525 | 3.571 | | | | | |
| 22 | 6.650 | 4.519 | 3.587 | 3.025 | | | | |
| 23 | 6.651 | 4.527 | 3.579 | 3.036 | | | | |
| 24 | 6.670 | 4.551 | 3.590 | 3.042 | | | | |
| 25 | 6.651 | 4.562 | 3.592 | 3.047 | | | | |
| 26 | 6.676 | 4.556 | 3.596 | 3.054 | | | | |
| 27 | 6.697 | 4.566 | 3.593 | 3.055 | | | | |
| 28 | 6.699 | 4.567 | 3.601 | 3.058 | | | | |
| 29 | 6.695 | 4.575 | 3.599 | 3.055 | | | | |
| 30 | 6.715 | 4.574 | 3.611 | 3.064 | | | | |
| 31 | 6.716 | 4.586 | 3.616 | 3.067 | | | | |
| 32 | 6.721 | 4.579 | 3.612 | 3.069 | 2.027 | | | |
| 33 | 6.728 | 4.593 | 3.617 | 3.075 | 2.036 | | | |
| 34 | 6.747 | 4.595 | 3.622 | 3.070 | 2.033 | | | |
| 35 | 6.740 | 4.596 | 3.616 | 3.076 | 2.038 | | | |
| 36 | 6.736 | 4.600 | 3.616 | 3.075 | 2.039 | | | |
| 37 | 6.729 | 4.610 | 3.621 | 3.081 | 2.034 | | | |
| 38 | 6.750 | 4.604 | 3.633 | 3.085 | 2.037 | | | |
| 39 | 6.745 | 4.617 | 3.641 | 3.085 | 2.051 | | | |
| 40 | 6.761 | 4.617 | 3.637 | 3.085 | 2.047 | | | |
| 41 | 6.747 | 4.624 | 3.634 | 3.086 | 2.052 | | | |
| 42 | 6.762 | 4.626 | 3.645 | 3.090 | 2.049 | 1.692 | | |
| 43 | 6.763 | 4.630 | 3.643 | 3.087 | 2.045 | 1.693 | | |
| 44 | 6.768 | 4.621 | 3.637 | 3.086 | 2.048 | 1.700 | | |
| 45 | 6.782 | 4.632 | 3.637 | 3.089 | 2.049 | 1.706 | | |
| 46 | 6.791 | 4.627 | 3.643 | 3.097 | 2.050 | 1.702 | | |
| 47 | 6.774 | 4.627 | 3.645 | 3.091 | 2.050 | 1.705 | | |
| 48 | 6.787 | 4.622 | 3.647 | 3.095 | 2.053 | 1.707 | | |
| 49 | 6.799 | 4.626 | 3.647 | 3.096 | 2.049 | 1.707 | | |
| 50 | 6.794 | 4.628 | 3.641 | 3.101 | 2.053 | 1.710 | | |
| 51 | 6.789 | 4.632 | 3.639 | 3.095 | 2.056 | 1.712 | | |
| 52 | 6.811 | 4.643 | 3.638 | 3.095 | 2.059 | 1.711 | 1.524 | |
| 53 | 6.797 | 4.636 | 3.640 | 3.092 | 2.050 | 1.715 | 1.527 | |
| 54 | 6.814 | 4.651 | 3.636 | 3.093 | 2.052 | 1.715 | 1.528 | |
| 55 | 6.796 | 4.637 | 3.647 | 3.089 | 2.061 | 1.714 | 1.531 | |
| 56 | 6.793 | 4.632 | 3.645 | 3.092 | 2.058 | 1.720 | 1.534 | |
| 57 | 6.811 | 4.637 | 3.647 | 3.093 | 2.059 | 1.718 | 1.535 | |
| 58 | 6.797 | 4.638 | 3.645 | 3.099 | 2.063 | 1.719 | 1.537 | |
| 59 | 6.804 | 4.635 | 3.651 | 3.100 | 2.061 | 1.716 | 1.538 | |
| 60 | 6.804 | 4.634 | 3.654 | 3.098 | 2.059 | 1.715 | 1.538 | |

| Control Limits $h_{n,\alpha=0.002,p}$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | $p$ | | | | |
| | 2 | 3 | 4 | 5 | 10 | 15 | 20 | 25 |
| $n$ | | | | | | | | |
| 61 | 6.801 | 4.629 | 3.656 | 3.091 | 2.063 | 1.721 | 1.540 | |
| 62 | 6.793 | 4.645 | 3.662 | 3.098 | 2.059 | 1.718 | 1.541 | 1.420 |
| 63 | 6.795 | 4.651 | 3.663 | 3.101 | 2.060 | 1.723 | 1.540 | 1.424 |
| 64 | 6.806 | 4.650 | 3.657 | 3.101 | 2.062 | 1.726 | 1.543 | 1.423 |
| 65 | 6.804 | 4.642 | 3.657 | 3.098 | 2.056 | 1.721 | 1.542 | 1.426 |
| 66 | 6.811 | 4.649 | 3.655 | 3.101 | 2.060 | 1.719 | 1.545 | 1.427 |
| 67 | 6.825 | 4.651 | 3.667 | 3.099 | 2.060 | 1.720 | 1.545 | 1.430 |
| 68 | 6.809 | 4.659 | 3.658 | 3.102 | 2.057 | 1.722 | 1.547 | 1.431 |
| 69 | 6.802 | 4.657 | 3.663 | 3.102 | 2.060 | 1.724 | 1.547 | 1.432 |
| 70 | 6.809 | 4.652 | 3.660 | 3.104 | 2.060 | 1.721 | 1.547 | 1.434 |
| 75 | 6.822 | 4.652 | 3.655 | 3.100 | 2.069 | 1.729 | 1.550 | 1.438 |
| 80 | 6.806 | 4.663 | 3.657 | 3.100 | 2.065 | 1.730 | 1.550 | 1.441 |
| 85 | 6.829 | 4.661 | 3.654 | 3.097 | 2.063 | 1.730 | 1.554 | 1.445 |
| 90 | 6.844 | 4.672 | 3.656 | 3.099 | 2.064 | 1.732 | 1.553 | 1.446 |
| 95 | 6.828 | 4.665 | 3.662 | 3.110 | 2.067 | 1.736 | 1.556 | 1.449 |
| 100 | 6.850 | 4.665 | 3.664 | 3.102 | 2.068 | 1.728 | 1.557 | 1.451 |
| 105 | 8.840 | 4.663 | 3.665 | 3.111 | 2.070 | 1.729 | 1.555 | 1.450 |
| 110 | 6.845 | 4.671 | 3.676 | 3.107 | 2.071 | 1.732 | 1.558 | 1.452 |
| 115 | 6.846 | 4.663 | 3.658 | 3.110 | 2.062 | 1.724 | 1.561 | 1.452 |
| 120 | 6.844 | 4.662 | 3.669 | 3.109 | 2.072 | 1.732 | 1.557 | 1.452 |
| 125 | 6.838 | 4.662 | 3.676 | 3.110 | 2.062 | 1.728 | 1.560 | 1.455 |
| 130 | 6.859 | 4.664 | 3.663 | 3.104 | 2.063 | 1.729 | 1.562 | 1.451 |
| 135 | 6.839 | 4.676 | 3.676 | 3.110 | 2.065 | 1.726 | 1.559 | 1.454 |
| 140 | 6.866 | 4.675 | 3.669 | 3.112 | 2.071 | 1.727 | 1.561 | 1.451 |
| 145 | 6.864 | 4.660 | 3.667 | 3.111 | 2.063 | 1.735 | 1.559 | 1.455 |
| 150 | 6.852 | 5.665 | 3.674 | 3.112 | 2.069 | 1.735 | 1.561 | 1.454 |

**Table 2.**

| | ARL when $\Sigma$ changes as $\begin{pmatrix} 1-\varrho & 0 \\ 0 & 1-\varrho \end{pmatrix}$ | | | | | |
|---|---|---|---|---|---|---|
| | | $\varrho$ | | | | |
| | | 0.1 | 0.2 | 0.4 | 0.6 | 0.8 |
| $n_0$ | $\delta$ | | | | | |
| 10 | 0.00 | 418.0 | 433.5 | 427.6 | 232.8 | 26.0 |
| | 0.25 | 382.2 | 399.2 | 400.4 | 191.7 | 22.9 |
| | 0.50 | 355.7 | 358.3 | 310.4 | 122.6 | 16.8 |
| | 1.00 | 124.6 | 118.9 | 72.6 | 23.5 | 10.9 |
| | 2.00 | 11.1 | 10.6 | 9.7 | 8.5 | 7.2 |
| | 3.00 | 6.9 | 6.7 | 6.6 | 6.4 | 6.0 |
| | 4.00 | 5.6 | 5.6 | 5.7 | 5.5 | 5.3 |
| 20 | 0.00 | 416.5 | 422.7 | 355.8 | 113.6 | 16.5 |
| | 0.25 | 387.9 | 376.4 | 318.1 | 93.9 | 16.2 |
| | 0.50 | 296.6 | 279.5 | 201.4 | 50.8 | 13.5 |
| | 1.00 | 68.1 | 50.5 | 27.5 | 16.1 | 10.3 |
| | 2.00 | 9.1 | 8.8 | 8.4 | 7.7 | 6.8 |
| | 3.00 | 6.1 | 6.0 | 5.9 | 5.8 | 5.6 |
| | 4.00 | 4.7 | 4.7 | 4.7 | 4.7 | 4.7 |
| 30 | 0.00 | 413.7 | 422.5 | 306.8 | 65.5 | 14.9 |
| | 0.25 | 369.1 | 370.1 | 255.2 | 48.5 | 14.5 |
| | 0.50 | 238.1 | 219.9 | 128.3 | 29.5 | 13.0 |
| | 1.00 | 32.4 | 30.4 | 21.2 | 14.3 | 9.8 |
| | 2.00 | 8.4 | 8.4 | 7.8 | 7.3 | 6.6 |
| | 3.00 | 5.7 | 5.7 | 5.6 | 5.4 | 5.2 |
| | 4.00 | 4.3 | 4.3 | 4.2 | 4.2 | 4.2 |
| 40 | 0.00 | 413.4 | 394.5 | 257.8 | 44.1 | 14.8 |
| | 0.25 | 359.3 | 340.1 | 200.0 | 39.3 | 14.4 |
| | 0.50 | 211.6 | 190.4 | 80.3 | 25.7 | 12.4 |
| | 1.00 | 25.9 | 25.6 | 19.7 | 14.5 | 9.6 |
| | 2.00 | 8.2 | 8.1 | 7.5 | 7.2 | 6.5 |
| | 3.00 | 5.4 | 5.3 | 5.3 | 5.2 | 5.1 |
| | 4.00 | 4.2 | 4.1 | 4.0 | 3.9 | 3.9 |

**Table 3.**

| | | ARL when $\Sigma$ changes as $\begin{pmatrix} 1 & \varrho \\ \varrho & 1 \end{pmatrix}$ | | | | |
|---|---|---|---|---|---|---|
| | | | | $\varrho$ | | |
| | | 0.1 | 0.2 | 0.4 | 0.6 | 0.8 |
| $n_0$ | $\delta$ | | | | | |
| 10 | 0.00 | 395.0 | 393.9 | 374.6 | 289.6 | 67.6 |
| | 0.25 | 366.4 | 368.8 | 358.9 | 245.0 | 66.1 |
| | 0.50 | 333.4 | 321.4 | 295.1 | 177.0 | 39.7 |
| | 1.00 | 144.6 | 142.3 | 117.3 | 50.4 | 18.5 |
| | 2.00 | 11.8 | 11.1 | 10.7 | 9.6 | 8.0 |
| | 3.00 | 6.9 | 6.8 | 6.8 | 6.6 | 6.4 |
| | 4.00 | 5.6 | 5.6 | 5.6 | 5.6 | 5.6 |
| 20 | 0.00 | 411.0 | 399.8 | 357.1 | 176.8 | 32.3 |
| | 0.25 | 368.1 | 377.4 | 313.0 | 159.5 | 31.6 |
| | 0.50 | 286.7 | 289.6 | 218.3 | 94.2 | 23.3 |
| | 1.00 | 69.1 | 57.9 | 43.5 | 24.6 | 13.1 |
| | 2.00 | 9.4 | 8.9 | 8.9 | 8.2 | 7.5 |
| | 3.00 | 6.1 | 6.0 | 6.0 | 6.0 | 5.8 |
| | 4.00 | 4.7 | 4.7 | 4.8 | 4.7 | 4.8 |
| 30 | 0.00 | 375.3 | 382.2 | 310.5 | 127.5 | 23.2 |
| | 0.25 | 381.5 | 362.0 | 260.1 | 111.1 | 22.0 |
| | 0.50 | 251.5 | 219.8 | 171.0 | 56.5 | 18.7 |
| | 1.00 | 38.5 | 33.9 | 28.2 | 18.0 | 12.4 |
| | 2.00 | 8.4 | 8.3 | 8.0 | 7.6 | 7.0 |
| | 3.00 | 5.6 | 5.5 | 5.6 | 5.4 | 5.4 |
| | 4.00 | 4.4 | 4.2 | 4.2 | 4.2 | 4.2 |
| 40 | 0.00 | 382.5 | 384.0 | 296.1 | 90.4 | 20.6 |
| | 0.25 | 351.1 | 312.5 | 214.1 | 71.8 | 20.3 |
| | 0.50 | 213.8 | 198.3 | 112.7 | 44.3 | 17.3 |
| | 1.00 | 27.1 | 26.8 | 22.9 | 16.8 | 11.8 |
| | 2.00 | 8.1 | 7.9 | 7.7 | 7.5 | 6.9 |
| | 3.00 | 5.5 | 5.3 | 5.3 | 5.3 | 5.2 |
| | 4.00 | 4.1 | 4.0 | 4.0 | 4.0 | 4.0 |

**Table 4.**

| | | ARL when $\Sigma$ changes as $\begin{pmatrix} 1+\varrho^2 & \varrho \\ \varrho & 1+\varrho^2 \end{pmatrix}$ | | | | |
|---|---|---|---|---|---|---|
| | | | | $\varrho$ | | |
| | | 0.1 | 0.2 | 0.4 | 0.6 | 0.8 |
| $n_0$ | $\delta$ | | | | | |
| 10 | 0.00 | 388.1 | 400.9 | 368.4 | 319.8 | 265.6 |
| | 0.25 | 373.5 | 367.9 | 337.9 | 312.4 | 259.1 |
| | 0.50 | 325.3 | 307.4 | 290.2 | 253.7 | 209.6 |
| | 1.00 | 144.7 | 140.8 | 132.0 | 109.3 | 98.4 |
| | 2.00 | 12.0 | 11.7 | 11.7 | 11.7 | 11.7 |
| | 3.00 | 7.0 | 6.9 | 6.9 | 6.9 | 7.0 |
| | 4.00 | 5.6 | 5.6 | 5.6 | 5.6 | 5.6 |
| 20 | 0.00 | 391.5 | 368.3 | 332.0 | 268.6 | 235.0 |
| | 0.25 | 347.8 | 341.4 | 316.8 | 254.6 | 200.5 |
| | 0.50 | 291.1 | 264.5 | 213.0 | 185.5 | 153.8 |
| | 1.00 | 73.2 | 70.2 | 51.6 | 43.8 | 37.2 |
| | 2.00 | 9.4 | 9.2 | 9.2 | 9.0 | 8.9 |
| | 3.00 | 6.1 | 6.1 | 6.1 | 6.0 | 6.0 |
| | 4.00 | 4.8 | 4.8 | 4.7 | 4.7 | 4.7 |
| 30 | 0.00 | 377.5 | 372.7 | 327.9 | 231.9 | 172.1 |
| | 0.25 | 363.3 | 332.8 | 309.0 | 208.5 | 158.6 |
| | 0.50 | 249.9 | 233.2 | 188.7 | 141.8 | 86.0 |
| | 1.00 | 39.9 | 37.7 | 32.2 | 25.6 | 21.8 |
| | 2.00 | 8.4 | 8.4 | 8.4 | 8.3 | 8.3 |
| | 3.00 | 5.6 | 5.6 | 5.7 | 5.6 | 5.5 |
| | 4.00 | 4.3 | 4.3 | 4.3 | 4.3 | 4.2 |
| 40 | 0.00 | 387.2 | 367.5 | 302.5 | 215.3 | 139.0 |
| | 0.25 | 353.4 | 331.7 | 245.1 | 169.4 | 109.4 |
| | 0.50 | 197.0 | 184.7 | 142.7 | 104.8 | 64.8 |
| | 1.00 | 28.2 | 26.7 | 24.0 | 21.5 | 19.6 |
| | 2.00 | 8.2 | 8.1 | 8.0 | 7.8 | 7.7 |
| | 3.00 | 5.4 | 5.4 | 5.4 | 5.4 | 5.3 |
| | 4.00 | 4.1 | 4.1 | 4.1 | 4.1 | 4.1 |

**Table 5.**

| Control Limits $h_{n,\alpha=0.002,p=2}$ | | | | | |
|---|---|---|---|---|---|
| $n$ | $h_{0,0.002,2}$ | $h_{10,0.002,2}$ | $h_{20,0.002,2}$ | $h_{30,0.002,2}$ | $h_{40,0.002,2}$ |
| 6 | 4.98 | | | | |
| 7 | 5.43 | | | | |
| 8 | 5.72 | | | | |
| 9 | 5.91 | | | | |
| 10 | 6.05 | | | | |
| 11 | 6.15 | | | | |
| 12 | 6.25 | | | | |
| 13 | 6.33 | | | | |
| 14 | 6.40 | | | | |
| 15 | 6.43 | | | | |
| 16 | 6.47 | 7.08 | | | |
| 17 | 6.50 | 6.79 | | | |
| 18 | 6.52 | 6.69 | | | |
| 19 | 6.55 | 6.65 | | | |
| 20 | 6.59 | 6.65 | | | |
| 21 | 6.59 | 6.64 | | | |
| 22 | 6.61 | 6.65 | | | |
| 23 | 6.62 | 6.65 | | | |
| 24 | 6.65 | 6.67 | | | |
| 25 | 6.64 | 6.65 | | | |
| 26 | 6.65 | 6.67 | 7.49 | | |
| 27 | 6.68 | 6.70 | 7.14 | | |
| 28 | 6.68 | 6.70 | 6.98 | | |
| 29 | 6.69 | 6.70 | 6.89 | | |
| 30 | 6.71 | 6.72 | 6.86 | | |
| 31 | 6.71 | 6.72 | 6.82 | | |
| 32 | 6.72 | 6.72 | 6.80 | | |
| 33 | 6.72 | 6.73 | 6.79 | | |
| 34 | 6.74 | 6.75 | 6.79 | | |
| 35 | 6.74 | 6.74 | 6.78 | | |
| 36 | 6.73 | 6.73 | 6.77 | 7.73 | |
| 37 | 6.72 | 6.73 | 6.76 | 7.30 | |
| 38 | 6.75 | 6.75 | 6.78 | 7.13 | |
| 39 | 6.74 | 6.75 | 6.77 | 7.02 | |
| 40 | 6.76 | 6.76 | 6.78 | 6.95 | |
| 41 | 6.74 | 6.75 | 6.76 | 6.90 | |
| 42 | 6.76 | 6.76 | 6.78 | 6.88 | |
| 43 | 6.76 | 6.76 | 6.78 | 6.86 | |
| 44 | 6.76 | 6.77 | 6.78 | 6.85 | |
| 45 | 6.78 | 6.78 | 6.79 | 6.84 | |
| 46 | 6.79 | 6.79 | 6.80 | 6.84 | 7.86 |
| 47 | 6.77 | 6.77 | 6.78 | 6.82 | 7.43 |
| 48 | 6.78 | 6.78 | 6.79 | 6.83 | 7.23 |
| 49 | 6.79 | 6.80 | 6.80 | 6.83 | 7.12 |
| 50 | 6.79 | 6.79 | 6.80 | 6.82 | 7.04 |
| 55 | 6.79 | 6.79 | 6.80 | 6.81 | 6.89 |
| 60 | 6.80 | 6.80 | 6.80 | 6.81 | 6.84 |
| 65 | 6.80 | 6.80 | 6.80 | 6.81 | 6.83 |
| 70 | 6.80 | 6.80 | 6.81 | 6.81 | 6.82 |
| 75 | 6.82 | 6.82 | 6.82 | 6.82 | 6.83 |
| 80 | 6.80 | 6.80 | 6.80 | 6.81 | 6.81 |
| 90 | 6.84 | 6.84 | 6.84 | 6.84 | 6.85 |
| 100 | 6.85 | 6.85 | 6.85 | 6.85 | 6.85 |
| 110 | 6.84 | 6.84 | 6.84 | 6.84 | 6.85 |
| 130 | 6.85 | 6.86 | 6.86 | 6.86 | 6.86 |
| 150 | 6.85 | 6.85 | 6.85 | 6.85 | 6.85 |

**Table 6.**

| Summary for change using step-down method | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $n_p$ | $\mu_{sys}$ | $\mu_{dia}$ | $\sigma_{sys}$ | $\rho_{sys,dia}$ | $\sigma_{dia}$ | $\chi^2_{in}$ | $\chi^2_{out\|in}$ | $p_{in}$ | $p_{out\|in}$ |
| 1 | 1 | 0 | 0 | 0 | 0 | 14.18 | 16.32 | 0.0001 | 0.0026 |
| 1 | 0 | 0 | 0 | 1 | 0 | 3.95 | 26.54 | 0.0467 | 0.0000 |
| 1 | 0 | 1 | 0 | 0 | 0 | 3.20 | 27.30 | 0.0735 | 0.0000 |
| 1 | 0 | 0 | 0 | 0 | 1 | 2.14 | 28.35 | 0.1426 | 0.0000 |
| 1 | 0 | 0 | 1 | 0 | 0 | 0.09 | 30.40 | 0.7574 | 0.0000 |
| 2 | 1 | 1 | 0 | 0 | 0 | 22.64 | 7.86 | 0.0000 | 0.0048 |
| 2 | 1 | 0 | 0 | 1 | 0 | 15.46 | 15.04 | 0.0004 | 0.0017 |
| 2 | 1 | 0 | 0 | 0 | 1 | 13.99 | 16.50 | 0.0009 | 0.0008 |
| 2 | 1 | 0 | 1 | 0 | 0 | 12.11 | 18.39 | 0.0023 | 0.0003 |
| 2 | 0 | 0 | 0 | 1 | 1 | 7.40 | 23.10 | 0.0246 | 0.0000 |
| 2 | 0 | 1 | 0 | 1 | 0 | 6.15 | 24.35 | 0.0462 | 0.0000 |
| 2 | 0 | 1 | 0 | 0 | 1 | 4.68 | 25.81 | 0.0959 | 0.0000 |
| 2 | 0 | 0 | 1 | 1 | 0 | 4.13 | 26.36 | 0.1262 | 0.0000 |
| 2 | 0 | 1 | 1 | 0 | 0 | 2.80 | 27.70 | 0.2463 | 0.0000 |
| 2 | 0 | 0 | 1 | 0 | 1 | 1.98 | 28.52 | 0.3716 | 0.0000 |
| 3 | 1 | 1 | 0 | 1 | 0 | 26.07 | 4.43 | 0.0000 | 0.1091 |
| 3 | 1 | 1 | 0 | 0 | 1 | 24.61 | 5.89 | 0.0000 | 0.0525 |
| 3 | 1 | 1 | 1 | 0 | 0 | 22.73 | 7.77 | 0.0000 | 0.0204 |
| 3 | 1 | 0 | 0 | 1 | 1 | 19.43 | 11.07 | 0.0002 | 0.0039 |
| 3 | 1 | 0 | 1 | 1 | 0 | 16.16 | 14.34 | 0.0010 | 0.0007 |
| 3 | 1 | 0 | 1 | 0 | 1 | 14.00 | 16.49 | 0.0028 | 0.0002 |
| 3 | 0 | 1 | 1 | 0 | 1 | 10.12 | 20.38 | 0.0175 | 0.0001 |
| 3 | 0 | 0 | 1 | 1 | 1 | 7.86 | 22.64 | 0.0489 | 0.0000 |
| 3 | 0 | 1 | 1 | 1 | 0 | 6.85 | 23.65 | 0.0767 | 0.0000 |
| 3 | 0 | 1 | 0 | 1 | 1 | 4.69 | 25.81 | 0.1956 | 0.0000 |
| 4 | 1 | 1 | 0 | 1 | 1 | 30.04 | 0.46 | 0.0000 | 0.4979 |
| 4 | 1 | 1 | 1 | 1 | 0 | 26.77 | 3.72 | 0.0000 | 0.0535 |
| 4 | 1 | 1 | 1 | 0 | 1 | 24.62 | 5.88 | 0.0000 | 0.0152 |
| 4 | 1 | 0 | 1 | 1 | 1 | 19.89 | 10.61 | 0.0000 | 0.0011 |
| 4 | 0 | 1 | 1 | 1 | 1 | 10.58 | 19.92 | 0.0000 | 0.0000 |

# References


Anderson T.W. (1984) An Introduction to Multivariate Statistical Analysis, John Wiley , New York.

Chambers, J. M.(1971), "Regression Updating". *Journal of the American Statistical Association.* 66, 744-748.

Champ, C. W., Jones-Farmer, L. A., and Rigdon, S. E., (2005)," Properties of the $T^2$ Control Chart when the Parameters are Estimated", *Tecnometrics*, 47, 437-445.

Croisier, R.B. (1988), " Multivariate generalizations of Cumulative Sum Quality-Control Schemes", *Technometrics,* 30, 291-303.

Hawkins, D. M., and Olwell, D. H., (1998), *Cumulative Sum Charts and Charting for Quality Improvement*, Springer Verlag, New York.

Hawkins, D. M., Qiu, P., Kang, C.W. (2003), "The Changepoint Model for Statistical Process Control" *Journal of Quality Technology*, 35, 355-365.

Hawkins, D. M., Zamba, K. D. (2005 a), " Change point Model for Statistical Process Control with Shift in Variance" *Journal of Quality Technology*, 37(1), 21-31.

Hawkins, D. M., Zamba, K. D. (2005 b), "Statistical Process Control for Shift in Mean or Variance Using the Change Point Formulation" *Technometrics* 47(2), 164-173.

Hawkins, D. M. Maboudou-Tchao E. M. (2007) "Self-Starting Multivariate Exponentially Weighted Moving Average Control Charting" *Technometrics* 49(2) 199-209(11).

Hawkins, D. M. Maboudou-Tchao E. M. (2008) "Multivariate Exponentially Weighted Moving Covariance Matrix" *Technometrics* 50(2), 155-166.

Healy, J. D.(1987), "A Note on Multivariate CUSUM Procedure", *Technometrics*, 29 409-412.

Hotelling, H (1947), " Multivariate Quality Control–Illustrated by the Air Testing of Bombsights", *Techniques of Statistical Analysis,* eds. C. Eisenhart, M.W.Hastay, and W.A. Wallis, New York: McGraw-Hill, pp 111-184.

Huwang, L., Yeh, A. B., and Wu, C.-W. (2007), "Monitoring Multivariate Process Variability for Individual Observations", *Journal of Quality Technology*, 39,258-278.

Lowry, C. A., Woodall W. H., Champ, C. W., and Rigdon S. E., (1992) " A Multivariate exponentially weighted moving average control chart ", *Technometrics*, 34, 46-53.

Margavio, T. M., Conerly, M. D., Woodall, W. H., and Drake, L. G., (1995), "Alarm Rates for Quality Control Charts", *Statistics and Probability Letters*, 24, 219-224.

Mason, R.L. and Young, J.C. (2002), *"Multivariate Statistical Process Control with Industrial Application"*, SIAM.

Pignatiello, J.J. and Runger, G.C. (1990), "Comparison of Multivariate Cusum Charts", *Journal of Quality Technology*, 22 (3), 173-186.

Reynolds, M. R., and Cho, G.-Y. (2006), "Multivariate Control Charts for Monitoring the Mean Vector and Covariance Matrix", *Journal of Quality Technology*, 38, 230-253.

Sullivan, J. H., Stoumbos, Z. G., Mason, R. L., and Young, J. C. (2007) "Step-Down Analysis for Changes in the Covariance Matrix and Other Parameters", *Journal of Quality Technology*, 39, 66-84.

Winkel P., Zhang N. F. ,(2007), Statistical Development of Quality in Medicine, John Wiley & Sons Ltd, England.

Woodall W. H., (2006),"The Use of Control Charts in Health-Care and Public-Health Surveillance",Journal of Quality technology, Vol. 38, No. 2, 89–104.

Zamba, K. D., and Hawkins, D. M. (2006) "A Multivariate Change Point Model for Statistical Process Control" *Technometrics* 48 , 4 ,539-549.