



Research Report  
Statistical Research Unit  
Department of Economics  
University of Gothenburg  
Sweden

---

## Evaluation of Multivariate Surveillance

**M. Frisé, E. Andersson &  
L. Schiöler**

**Research Report  
2009:1  
ISSN 0349-8034**

---

Mailing address:	Fax	Phone	Home Page:
Statistical Research Unit	Nat: 031-786 12 74	Nat: 031-786 00 00	<a href="http://www.statistics.gu.se/">http://www.statistics.gu.se/</a>
P.O. Box 640	Int: +46 31 786 12 74	Int: +46 31 786 00 00	
SE 405 30 Göteborg			
Sweden			

# Evaluation of multivariate surveillance

Marianne Frisén<sup>a\*</sup>, Eva Andersson<sup>a,b</sup> and Linus Schiöler<sup>a</sup>

*<sup>a</sup>Statistical Research Unit, Department of Economics, University of Gothenburg*

*<sup>b</sup>Department of Occupational and Environmental medicine, Sahlgrenska University Hospital and Sahlgrenska Academy, University of Gothenburg*

Multivariate surveillance is of interest in many areas such as industrial production, bioterrorism detection, spatial surveillance, and financial transaction strategies. Some of the suggested approaches to multivariate surveillance have been multivariate counterparts to the univariate Shewhart, EWMA, and CUSUM methods. Our emphasis is on the special challenges of evaluating multivariate surveillance methods. Some new measures are suggested and the properties of several measures are demonstrated by applications to various situations. It is demonstrated that zero-state and steady-state ARL, which are widely used in univariate surveillance, should be used with care in multivariate surveillance.

Keywords: average run length; delay; EWMA; false alarms; FDR; performance metrics; predictive value; steady state; zero state

## 1. Introduction

In many situations there are reasons to continuously observe a process in order to detect an important change in the process as soon as possible after the change has occurred. Multivariate surveillance typically concern several variables. However it is also of interest when there is only one process, but several characteristics of that process may change. Examples are processes where both the mean and the variance may change (treated in [19]) or changes in several aspects of one autoregressive time series as treated in [3].

The first suggestion of modern control charts [33] was widely utilized by industry. The monitoring of several processes is often of interest. Multivariate problems for the assembly process of the Saab automobile were described in [40]. In food industry different raw materials and several process steps are used, and in [32] it is suggested that these be analyzed in order to assure the quality of the final product. During the last years there have been an increased need for and interest in continuous monitoring in many areas apart from industrial production. After the 9/11 attack the interest in surveillance methodology increased notably in the US, and new types of data are now being collected to get early signals of bioterrorism. By monitoring several data series different aspects can be covered, and thus multivariate surveillance techniques are needed. [28] contains an overview of the research needs for bioterrorism surveillance using multiple data streams. Spatial surveillance is another example of multivariate surveillance, since several locations are involved. A relatively new area for multivariate surveillance is financial decision strategies in situations where a portfolio contains several assets (see for example [13, 25]).

General reviews on multivariate surveillance methods are made for example in [4, 6, 9, 21, 31, 36, 39].

Multivariate surveillance can have different aims. Sometimes, the aim is to identify which parameters that have changed (e.g. identify faulty components). However, this is naturally preceded by the detection of a change in any of the parameters. Here we concentrate on the detection of the first change.

---

\*Corresponding author. Email: marianne.frisen@statistics.gu.se

## 2. Notations and specifications

We will denote the multivariate process under surveillance by a  $p$ -variate vector,  $\mathbf{Y}(t) = \{Y_1(t), Y_2(t), \dots, Y_p(t)\}$ . The components of the vector may be, for example, a measurement on  $p$  different components. The distribution of the  $p$ -variate variable  $\mathbf{Y}(t)$  might be characterized by the mean vector  $\boldsymbol{\mu}$  and covariance matrix  $\boldsymbol{\Sigma}_Y$ . The aim is to detect the change from one state – for example that the assembled product works well – to another – that some component is defective so that the product does not work. We aim to detect the change as soon as possible after it has occurred in order give warnings and to take corrective actions. At decision time  $s$  we base the decision on the available information  $\mathbf{Y}^s = \{\mathbf{Y}(1), \mathbf{Y}(2) \dots \mathbf{Y}(s)\}$  to form an alarm statistic. An alarm is called the first time that the statistic exceeds an alarm limit,  $G(s)$ .

In the multivariate situation we observe  $p$  processes which can change at different times  $\tau_1, \dots, \tau_p$ . Here the aim is often to detect the first time that the process is no longer in control – that is, we want to make inference about  $\tau_{\min} = \min\{\tau_1, \dots, \tau_p\}$ . If there is no change at all, we denote this by “ $\tau_{\min} = \infty$ ”.

## 3. Surveillance methods

We will discuss different evaluation metrics for multivariate surveillance. The discussion is supported by results where commonly used methods are evaluated by the metrics. The evaluation measures will reveal principal differences between approaches for multivariate surveillance.

### 3.1. General approaches

#### 3.1.1. Dimension reduction

One approach for handling multivariate surveillance problems is to reduce the  $p$ -variate vector at each time point into a single statistic and then use a system for univariate surveillance on this statistic. One may simply use the sum or another linear combination of the variables. When we want to derive an optimal method, we must specify the type of change that we want the method to detect. One way to focus the attention is to consider some type of dimension reduction transformation as in [14, 15, 29]. In [30] this is done with specific respect to the special and common causes of variation. Sometimes a sufficient reduction can be found as in [37] where it is proved that when the changes occur simultaneously, it is possible to find a sufficient reduction to a univariate surveillance problem. For the exponential family with the same shift and dispersion parameter and independence between the processes, conditional on change times, the sufficient statistic for each time  $t$  is the sum of the observations  $Y_1(t), \dots, Y_p(t)$ . For some situations where the changes occur with known lags, it is also possible to find a sufficient reduction, see [17]. When a sufficient reduction is found, optimal methods can be derived. In many situations, however, it is not possible to find a sufficient reduction.

#### 3.1.2. Parallel surveillance

The approach with parallel systems means that one starts with a univariate surveillance method for each variable. The most common way to combine the information from the  $p$  univariate methods is to signal an alarm if any of the univariate methods gives an alarm. This approach is called combined univariate methods or parallel methods.

### 3.2. Specific methods and situations

#### 3.2.1. Example

We will illustrate the suggested measures and their properties by applying them in a number of different situations and for different methods. We will concentrate on the way in which the time of the changes influences the properties, and therefore a very simple example with two processes will be used. Our model contains two normally distributed variables,  $Y_1$  and  $Y_2$ , which possibly have shifts in the expected value at possibly different time points. In order to focus on the effect of different change times we use equal shift sizes. The two processes,  $Y_1$  and  $Y_2$ , are assumed to be independent (conditional on the change times).

$$Y_1(t) \sim \begin{cases} N(0,1) & t < \tau_1 \\ N(1,1) & t \geq \tau_1 \end{cases}$$

$$Y_2(t) \sim \begin{cases} N(0,1) & t < \tau_2 \\ N(1,1) & t \geq \tau_2 \end{cases}$$

The alarm times for different methods were determined by Monte Carlo simulations with at least 10,000,000 replicates for each situation.

#### 3.2.2. Specific methods

Multivariate methods are usually extensions of common univariate methods. The univariate technique used here is the EWMA method, since it is commonly used also in multivariate situations. As regards the variance of the EWMA statistic there are two versions: the exact and the asymptotic variance, and we will use the asymptotic version as recommended in [35]. The statistic of the EWMA method for univariate surveillance of a variable  $Y$  is

$$Z_s = \lambda(1-\lambda)^s \sum_{t=1}^s (1-\lambda)^{-t} Y(t),$$

where  $0 < \lambda \leq 1$  and  $Z_0$  is the target value, which here is zero.

The optimal value of the parameter  $\lambda$  has drawn much attention. A formula for the optimal value was derived in [9] and explicitly given in [11] as  $\lambda^* = 1 - \exp(-\mu^2/2)/(1-\nu)$ , where  $\mu$  is the shift size (here  $\mu = 1$ ) and  $\nu = P(\tau=t \mid \tau \geq t)$  denotes how often changes are prone to occur. Here we choose the value  $\lambda=0.35$  which will give an approximately optimal method for a wide range of  $\nu$ .

We will compare results from i) the EWMA method applied to a reduction of data to a univariate statistic at each time, ii) a system based on two parallel EMWA methods, and iii) the EWMA method applied to the univariate process that changes first.

As an example of the reduction approach, we reduce the bivariate variable  $(Y_1, Y_2)$  to a univariate statistic, here chosen to

$$R(t) = (Y_1(t) + Y_2(t))/2.$$

Then the EWMA method is applied to the variable  $R(t)$  (with the variance  $\sigma_R^2 = 0.5$ ). The time of alarm for the reduction method,  $t_{AR}$ , is the first time when the EWMA statistic exceeds a constant alarm limit.

The parallel approach means that the EWMA method is applied to  $Y_1(t)$  and  $Y_2(t)$  separately. The time of alarm for  $Y_1$ ,  $t_{A1}$ , is the first time when  $EWMA_{Y_1}$  exceeds a constant alarm limit (correspondingly for  $Y_2$ ). The time of alarm for the parallel approach is the first of either of the alarm times ( $t_{AP} = \min[t_{A1}, t_{A2}]$ ).

For comparison we also have the results from the EWMA method applied to only one process. This corresponds to the situation when there is prior knowledge about which process will change first and therefore it is efficient to monitor only this one.

The alarm limits are set in order to give each of the systems the same false alarm property.

## 4. Evaluation metrics

The timeliness in detection is of extreme interest in surveillance, and hence there is a need for other evaluation measures than the ones traditionally used in hypothesis testing.

### 4.1. False alarms

In a univariate setting the most commonly used measure is  $ARL^0 = E[t_A | \tau = \infty]$ . This is naturally generalized as  $E[t_A | \tau_{\min} = \infty] = E[t_A | \tau_1 = \infty, \dots, \tau_p = \infty]$ . The median run length,  $MRL^0$ , can be used instead of the expected value with the same generalization as for  $ARL^0$ . In the simulations below, the alarm limits are set so that each of the systems has an  $MRL^0$  equal to 100.

In theoretical work the false alarm probability,  $PFA = P(t_A < \tau)$ , is commonly used. This is naturally generalized as

$$PFA = P(t_A < \tau_{\min}) = \sum_{i=1}^{\infty} P(t_A < \tau_{\min} | \tau_{\min} = i) P(\tau_{\min} = i)$$

It can also be expressed as  $PFA = P(t_A < \tau_j) P(\tau_{\min} = \tau_j)$

Note that the distribution of  $\tau_{\min}$  (through the distribution of the change point distributions of all variables) is included in the suggested multivariate PFA expression.

In hypothesis testing with multiple comparisons it is important to control the probability of false rejection (an overview of important methods is given in [16]). For the situation when several drugs are tested against one standard, the family-wise error rate is relevant. For another situation, for example when several aspects of a single drug are tested, the False Discovery Rate (FDR), suggested in [5], may be more relevant. Recently FDR has been suggested for surveillance problems for example in [28]. Surveillance, where we make more than one decision, differs from hypothesis testing in that methods with high detection ability have a false alarm rate that tends to one (as time tends to infinity), see for example [7]. If one tries to avoid this, by letting the alarm limit tend to infinity, it will harm the ability to detect late changes. Thus, false alarms are not regarded in the same way in surveillance as in hypothesis testing. The FDR measure is difficult to use in surveillance, since it is based on a probability which is not constant. There are different suggestions for solving this problem: In [23] a fairly short period of time is monitored and only the properties of the early part of the run length is used. When surveillance is used as a screening instrument, with follow-up tests, it may be less important to control the FDR. The  $ARL^0$  of the multivariate procedure, as suggested above, might be easily interpreted as the expected time until an un-necessary screening.

### 4.2. Delay

#### 4.2.1. Delay as a function of the time of the change

We start by recapturing the univariate case where the expected delay for a specific value of  $\tau$  is

$$ED(\tau) = E\{(t_A - \tau)^+\},$$

or, if  $\tau$  is stochastic, the average delay over the distribution of  $\tau$

$$ED = E\{ED(t)\}.$$

This average is the base for the ED optimality, which is closely related to the utility functions suggested in [34] and sometimes called a Bayesian measure since it depends on the distribution of  $\tau$ , which for some applications is naturally regarded as a parameter and for others as a stochastic variable.

Since  $ED(\tau)$  for most methods tends to zero (because of the false alarms when  $\tau$  tends to infinity), it is useful to study the delay conditional on no alarms before  $\tau$ . For a specific value of  $\tau$ , the Conditional Expected Delay, CED, is

$$CED(\tau) = E[t_A - \tau | t_A \geq \tau].$$

The first use of the term CED and calculation for a specific value of  $\tau$ , different from 1 and  $\infty$  seems to be in [41]. In [1, 12], the CED was used as a function of  $\tau$ , and in [9, 11] it was strongly advocated that the whole CED curve be studied. In [20] the dependency on  $\tau$  is avoided by using the least favorable value of  $\tau$ . The asymptotic measure is another example of how the value of  $\tau$  can be avoided. The CED has been a component in many measures but often in a way which avoids the dependency on  $\tau$ .

In the multivariate case the  $ED(\tau_1, \dots, \tau_p)$  and  $CED(\tau_1, \dots, \tau_p)$  depend on the vector  $\{\tau_1, \dots, \tau_p\}$ , and ED depends on the multivariate distribution of  $(\tau_1, \dots, \tau_p)$ . [2] suggested the following delay measure (for a situation where  $p=2$ )

$$CED(\tau_1, \tau_2, \dots, \tau_p) = E[t_A - \tau_{\min} | \tau_{\min} \geq t_A]$$

and demonstrated the dependency on  $\tau_{\min}$ . This delay measure depends on all the change points. However, there is often some relation between the change times which simplifies the picture. In Figure 1 and 2, we will use the multivariate CED to demonstrate principal differences between methods for some typical situations with special relations between the change times. In Figure 1 the conditional expected delay is presented for the Parallel and Reduction approaches, for the example where the changes appear simultaneously.

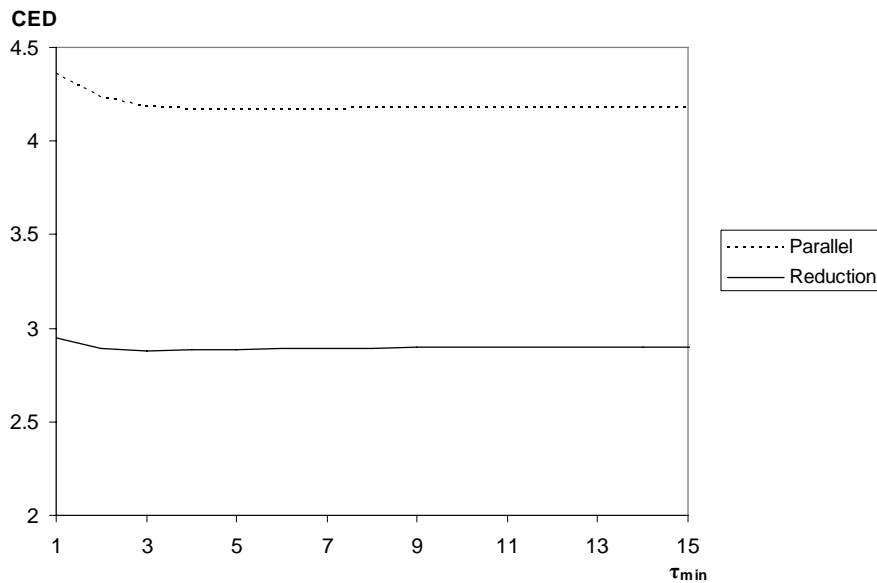


Figure 1.  $CED(\tau_1, \tau_2)$  vs  $\tau_{\min}$  for  $\tau_1=\tau_2=\tau_{\min}$ , presented for the Parallel and Reduction methods.

From Figure 1 we can see that the results in [37] mentioned in Section 3.1.1 hold here: the Reduction method is the best (gives the shortest delay) when all processes change at the same time. In Figure 2 it is seen that the CED curves differ considerably for different relations between the values of the change times.

Sometimes the time available for action is limited. In such situations it is important to use a surveillance system with high detection ability within the limited time available. This property can be measured by the Probability of Successful Detection, which was suggested by [8]. It measures the probability that an alarm is called within  $d$  time points. In the multivariate case it can be defined as

$$PSD(d, \tau_1 \dots \tau_p) = P(t_A - \tau_{\min} \leq d | t_A \geq \tau_{\min}),$$

as in [10].

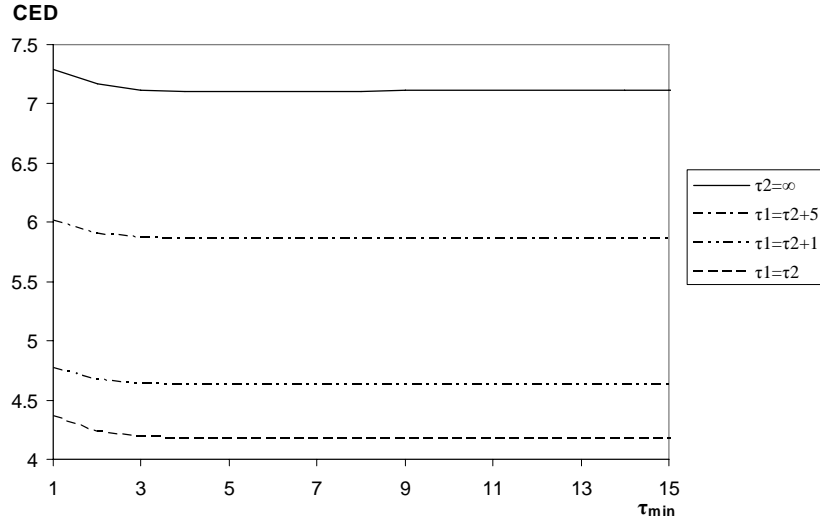


Figure 2.  $CED(\tau_1, \tau_2)$  vs  $\tau_{\min}$  for different relations between the  $\tau$  values, presented for the Parallel method.

The PSD measure is a function of both the change times ( $\tau_1, \dots, \tau_p$ ) and the length of the interval in which the detection is defined as successful ( $d$ ). [38] suggested that the PSD be calculated as a function of only  $d$  and  $\tau_{\min}$ , by expressing PSD as an expected value for other (stochastic) change points than  $\tau_{\min}$ . The PSD can be used to describe the detection ability of a method and compare it to that of other methods. PSD can also be calculated and compared for different values of  $d$ , as is done in [23] in connection with the use of the FDR (false detection rate). If we expect sudden and major changes, we may want a method with high detection ability (a high PSD for a small  $d$ ). In a situation where we expect small changes, the long term detection ability (a high PSD for a large  $d$ ) may be more important. Thus it is essential to consider what kind of change one wants to detect at different time points. In Figure 3 we examine the PSD for the Parallel method, for two different cases of relations between the change points. With the Parallel method, it is easier to quickly detect simultaneous changes than changes with a time lag. The PSD will tend to one for both cases when  $d$  increases.

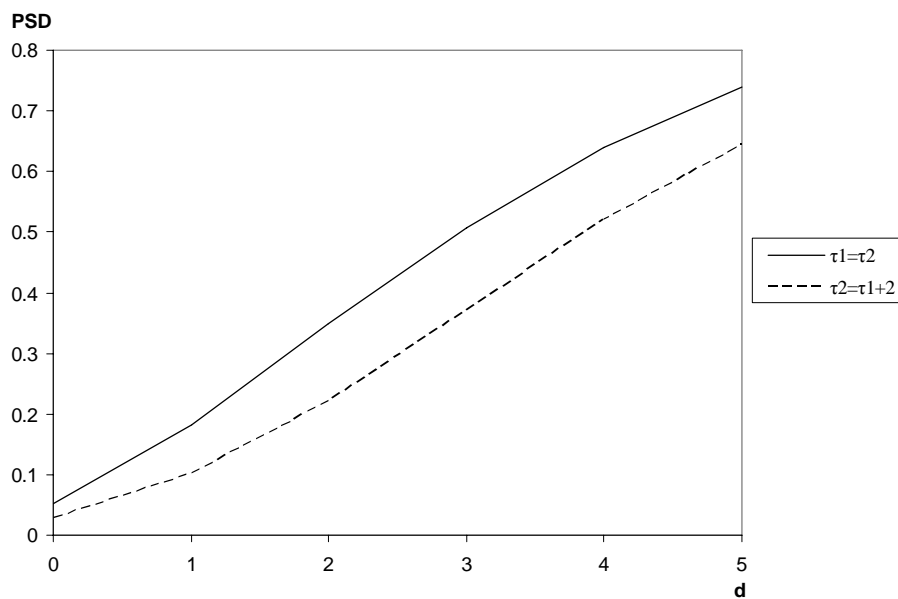


Figure 3.  $PSD(\tau_1, \tau_2)$  vs  $d$  for different relations between the  $\tau$  values ( $\tau_1 = \tau_2$  and  $\tau_1 + 2 = \tau_2$ ) presented for the Parallel method when  $\tau_{\min} = 3$ .

#### 4.2.2. Zero-state ARL

One measure of the detection ability is the average run length, given that the change occurs immediately ( $\tau=1$ ). This is widely used in univariate surveillance and often named zero-state ARL or  $ARL^1$ . In univariate surveillance the  $ARL^1$  has a simple relation to the delay, namely  $ARL^1=CED(1)+1$ . This demonstrates that only  $\tau=1$  is considered. It may also be important to consider other change times, since the delay and detection ability of many methods depend on when the change occurs (i.e. depend on  $\tau$ ). To consider only  $\tau=1$  in the univariate case is a limitation, and the univariate  $ARL^1$  is criticized as a formal optimality criterion for example by [9].

Zero-state ARL is the most commonly used evaluation measure also in the multivariate case. However, it is seldom explicitly defined. One possibility is to define the multivariate zero-state ARL as  $E[t_A|\tau_{\min}=1]$ . However, as seen in Figure 2, the values of CED for  $\tau_{\min}=1$  vary a lot for different relations between the values of  $\tau_{\min}$  and the change times of the other processes. Thus, there is no unique zero-state ARL with the definition  $E[t_A|\tau_{\min}=1]$ . Another possibility is to define the multivariate zero-state ARL as  $E[t_A|\tau_1=\tau_2=\dots=\tau_p=1]$ . This is probably the definition implicit in most publications. Here, it is assumed that all processes change at the same time. It was demonstrated by [37] that a sufficient reduction to a univariate problem exists when all processes change at the same time. Thus, when  $\tau_1=\dots=\tau_p=1$ , a reduction to a univariate surveillance statistic is the proper procedure by the sufficiency principle, which means that we have a univariate situation. Zero-state ARL is thus questionable as a formal measure for comparing methods for genuinely multivariate problems.

#### 4.2.3. Steady-state delay

Already [27] suggested the use of the limit of CED as  $\tau$  tends to infinity (even though he used  $\tau=8$  in the numerical comparisons). Here this will be called the steady-state conditional expected delay,  $CED_{SS}$ .

$$CED_{SS} = \lim_{\tau \rightarrow \infty} CED(\tau)$$

This steady-state delay is closely related to steady-state ARL (often denoted by SS ARL or  $ARL_{SS}$ ) which is defined in [22] as “The time from the change to the signal ... using the steady state distribution” or more specifically by [18] as

$$\lim_{\tau \rightarrow \infty} E[t_A - \tau + 1 | t_A \geq \tau].$$

Here we see that  $ARL_{SS}=CED_{SS} + 1$ . This corresponds to the relation  $ARL^1=CED(1)+1$ .

Evaluations of multivariate methods by asymptotic measures are often made by the same measures as are used for univariate methods. For example, [22] used “the steady state average delay time” and [26] “the steady-state Average Time to Signal.” However, the correspondence to the univariate  $CED_{SS}$  is not without problems. The multivariate CED depends on several  $\tau$  values and so does the multivariate steady-state conditional expected delay, as seen in Figure 2. There is thus no unique steady-state CED (or steady-state ARL) that could characterize a method. Often only the situation  $\tau_1=\tau_2=\dots=\tau_p$  is considered. In that case we have

$$CED(\tau_1=\tau_2=\dots=\tau_p=t) \text{ as } t \rightarrow \infty.$$

For equal change points we have a unique delay value for each method. However, this is another example of the situation where univariate surveillance can be used instead of multivariate surveillance since there is a sufficient reduction to univariate surveillance. This is confirmed in Figure 1, where we saw that the best method is based on the reduction to a univariate statistic. For other situations than simultaneous changes there is no simple asymptotic CED, as is seen in Figure 2. Even though all the  $\tau$  values tend to infinity, it also matters how they do this. There is no simple asymptotic measure for the multivariate case. Instead, one has to specify how the times of the change points are related when they tend to infinity.



### 4.3. Predictive value

The predictive value, suggested by [8], is defined as

$$PV(t) = P(C(t) | t_A = t) = \frac{PMA(t)}{PMA(t) + PFA(t)},$$

where PMA is the probability of a motivated alarm and PFA is the probability of a false alarm. Thus,  $PMA(t) = P(C(t) | t_A = t)$  and  $PFA = P(D | t_A = t)$ .

In a univariate setting with  $C(t) = \{\tau \leq t\}$  and  $D = \{\tau > t\}$  this is

$$PV(t) = P(\tau \leq t | t_A = t) = \frac{\sum_{i=1}^t P(t_A = t | \tau = i) \cdot P(\tau = i)}{\sum_{i=1}^t P(t_A = t | \tau = i) \cdot P(\tau = i) + P(t_A = t | \tau > t) \cdot P(\tau > t)}$$

In a multivariate setting we generalize this with  $C(t) = \{\tau_{\min} \leq t\}$  and  $D = \{\tau_{\min} > t\}$  to

$$PV(t) = P(\tau_{\min} \leq t | t_A = t) = \frac{\sum_{i=1}^t P(t_A = t | \tau_{\min} = i) \cdot P(\tau_{\min} = i)}{\sum_{i=1}^t P(t_A = t | \tau_{\min} = i) \cdot P(\tau_{\min} = i) + P(t_A = t | \tau_{\min} > t) \cdot P(\tau_{\min} > t)}$$

For the case of two variables,  $Y_1$  and  $Y_2$ , we have that the probabilities of a motivated and a false alarm, respectively, are

$$PMA(t) = \sum_{i=1}^t \sum_{j=1}^t (P(t_A = t | \tau_1 = i, \tau_2 = j) P(\tau_1 = i, \tau_2 = j)) + \sum_{i=1}^t (P(t_A = t | \tau_1 = i, \tau_2 > t) P(\tau_1 = i, \tau_2 > t)) + \sum_{j=1}^t (P(t_A = t | \tau_1 > t, \tau_2 = j) P(\tau_1 > t, \tau_2 = j))$$

and

$$PFA(t) = P(t_A = t | \tau_1 > t, \tau_2 > t) \cdot P(\tau_1 > t, \tau_2 > t)$$

For independently geometrically distributed change processes with the same intensity  $\nu$ , the alarm probabilities simplify. If also the distributions between which the changes appear are the same for the two variables as in the example, we get

$$PMA(t) = \sum_{i=1}^t \sum_{j=1}^t (P(t_A = t | \tau_1 = i, \tau_2 = j) [\nu^2 (1-\nu)^{i+j-2}]) + 2 \sum_{i=1}^t (P(t_A = t | \tau_1 = i, \tau_2 > t) [\nu (1-\nu)^{i+t-1}])$$

and

$$PFA(t) = P(t_A = t | \tau_1 > t, \tau_2 > t) \cdot (1-\nu)^{2t}$$

In Figure 4, the predicted value is illustrated for two methods, Parallel and Reduction. We can see that the Parallel method has a better PV than the Reduction method. This can be expected since the change points are seldom simultaneous when we have independent processes with low intensities.

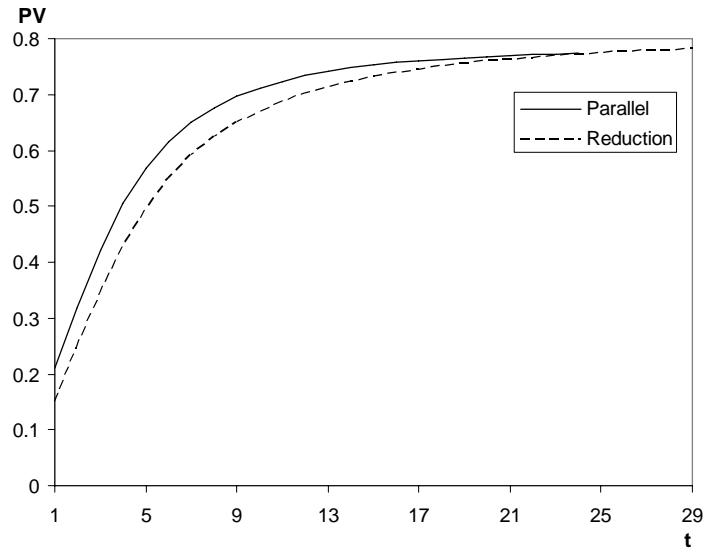


Figure 4. The predicted value (for  $C(t) = \{\tau_{\min} \leq t\}$ ) at different alarm times  $t_A$ , for the case where  $\tau_1$  and  $\tau_2$  are independently geometrically distributed with parameter values  $\nu_1 = \nu_2 = 0.01$ .

For simultaneous changes with  $\tau_1 = \tau_2 = \tau$  and  $\tau$  geometrically distributed with  $\nu = 0.01$ , we have that the probabilities of a motivated and a false alarm, respectively, are

$$PMA = \sum_{i=1}^t (P(t_A = t | \tau = i) [\nu(1-\nu)]^{i-1})$$

and

$$PFA(t) = P(t_A = t | \tau > t) \cdot (1-\nu)^t.$$

As seen in Figure 5, the Reduction method has a better predicted value than the Parallel method when both processes change at the same time. By comparing Figure 4 and 5, we see that the method which has the best predicted value and thus the most trustworthy alarms depends on the relation between the change points.

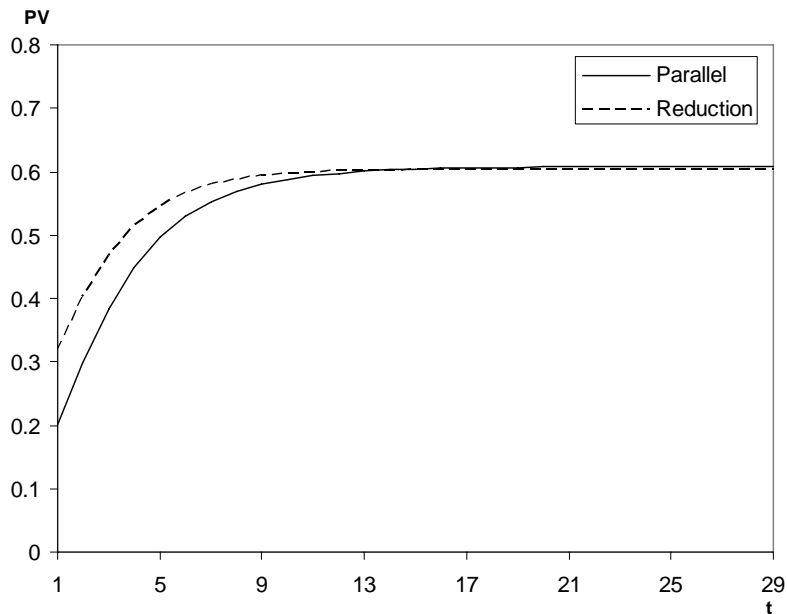


Figure 5. The predicted value at different alarm times  $t_A$ , for the case where  $\tau_1 = \tau_2 = \tau$  and  $\tau$  geometrically distributed with  $\nu = 0.01$ .

## 5. Discussion

Optimality is often hard to define in multivariate problems due to the several dimensions resulting from the variables. A method could work well for detecting a change in one direction but not in others. In surveillance (univariate as well as multivariate), evaluation is difficult due to the complex time relations. Some methods work well for detecting gradual long term changes and others for detecting sudden large ones. Thus, it is a challenge to evaluate multivariate surveillance methods which involve the difficulties with both several dimensions and complex time relations. The use of multivariate surveillance methods is growing, and the evaluation challenge has to be approached.

Some new measures, which are generalizations of univariate counterparts, were suggested here and the properties of several measures were demonstrated by applications to various situations. The relation between the change times is very important for deciding which method is the best. For example, the Reduction method gives the shortest delay and the highest predictive value when all processes change at the same time but not when the changes occur separately. The Parallel approach has a higher predictive value when the changes are not prone to occur simultaneously.

It was demonstrated that zero-state and steady-state ARL, which are widely used in univariate surveillance, should be used with care in multivariate surveillance. Unfortunately, the more elaborated CED measure is necessary for full information.

The numerical values of the evaluation measures can be hard to obtain analytically for surveillance methods. Thus, Monte Carlo simulations (as in this paper and many others) or numerical approximations (as for example by [18]) are useful. Evaluation by application to a single case might be interesting but has the drawback of being highly dependent on stochastic variation. Applications to several cases diminish this drawback. An approach between the application to a single case and simulations is the technique of using an observed data series as a start and inducing simulated disturbances to this series (see for example [24]).

For the measures PFA, ED, and PV, we need the distribution of  $\tau_{\min}$ , which in turn depends on the distributions of the change times for all processes. These measures are only suitable when the change process is considered to be stochastic. The other measures are also suitable when the change points are considered as unknown but fixed values.

Even if it is appropriate for the application to consider the change points as stochastic, the exact distribution is seldom known. However, any indication about the predicted value is of great importance for the interpretation of an alarm. An alarm does not give cause for extensive action if the predicted value is low. In Figure 4 we can see that the predictive value can be low for early alarms. This means that these should not call for the same actions as later alarms.

## Acknowledgements

Kjell Pettersson has given constructive comments throughout the work. The work was supported by the Swedish Emergency Management Agency (grant 0314/206). The authors have declared no conflict of interest.

## References

- [1] E. Andersson, *Monitoring cyclical processes - A nonparametric approach*, Journal of Applied Statistics 29 (2002), pp. 973-990.
- [2] ---, *Effect of dependency in systems for multivariate surveillance*, Communications in Statistics. Simulation and Computation 38 (2009), pp. 454-472.
- [3] D.W. Apley, and F. Tsung, *The Autoregressive T-squared Chart for Monitoring Univariate Autocorrelated Processes*, Journal of Quality Technology 34 (2002), pp. 80-96.
- [4] M. Basseville, and I. Nikiforov, *Detection of Abrupt changes - Theory and Application*, ed, Prentice Hall, Englewood Cliffs, 1993.

- [5] Y. Benjamini, and Y. Hochberg, *Controlling the false discovery rate: a practical and powerful approach to multiple testing*, Journal of the Royal Statistical Society, Series B 57 (1995), pp. 289-300.
- [6] S. Bersimis, S. Psarakis, and J. Panaretos, *Multivariate Statistical Process Control Charts: An Overview*, *Quality and Reliability Engineering International* 23 (2007), pp. 517-543.
- [7] D. Bock, *Aspects on the control of false alarms in statistical surveillance and the impact on the return of financial decision systems*, Journal of Applied Statistics 35 (2008), pp. 213 - 227.
- [8] M. Frisé, *Evaluations of Methods for Statistical Surveillance*, *Statistics in Medicine* 11 (1992), pp. 1489-1502.
- [9] ---, *Statistical surveillance. Optimality and methods*, *International Statistical Review* 71 (2003), pp. 403-434.
- [10] ---, *Principles for Multivariate Surveillance*, to appear in *Frontiers in Statistical Quality Control*, H.-J. Lenz and P.-T. Wilrich eds., 2009.
- [11] M. Frisé, and C. Sonesson, *Optimal surveillance based on exponentially weighted moving averages*, *Sequential Analysis* 25 (2006), pp. 379-403.
- [12] M. Frisé, and P. Wessman, *Quality improvement by likelihood ratio methods for surveillance.*, in *Quality Improvement Through Statistical Methods*, B. Abraham ed., Birkhauser, Boston, 1998, pp. 187-193.
- [13] V. Golosnoy, W. Schmid, and I. Yatsyshynets, *Sequential Monitoring of Optimal Portfolio Weights*, in *Financial surveillance*, M. Frisé ed., Wiley, Chichester, 2007, pp. 179-210.
- [14] S. Hao, S. Zhou, and Y. Ding, *Multivariate Process Variability Monitoring Through Projection* *Journal of Quality Technology* 40 (2008), pp. 214-226.
- [15] D.M. Hawkins, *Multivariate Quality Control Based on Regression-Adjusted Variables*, *Technometrics* 33 (1991), pp. 61-75.
- [16] Y. Hochberg, and A.C. Tamhane, *Multiple comparison procedures*, ed, Wiley, New York, 1987.
- [17] E. Järpe, *On univariate and spatial surveillance. Ph.D Thesis*, Göteborg University, 2000.
- [18] S. Knoth (ed.), *The art of evaluating monitoring schemes - how to measure the performance of control charts?*, ed, Physica Verlag, Warsaw, 2006.
- [19] S. Knoth, and W. Schmid, *Monitoring the mean and the variance of a stationary process*, *Statistica Neerlandica* 56 (2002), pp. 77-100.
- [20] G. Lorden, *Procedures for reacting to a change in distribution*, *Annals of Mathematical Statistics* 42 (1971), pp. 1897-1908.
- [21] C.A. Lowry, and D.C. Montgomery, *A Review of Multivariate Control Charts*, *IIE Transactions* 27 (1995), pp. 800-810.
- [22] C.W. Lu, and M.R. Reynolds Jr, *EWMA control charts for monitoring the mean of autocorrelated processes*, *Journal of Quality Technology* 31 (1999), pp. 166-188.
- [23] C. Marshall et al., *Statistical issues in the prospective monitoring of health outcomes across multiple units*, *Journal of the Royal Statistical Society A* 167 (2004), pp. 541-559.
- [24] M. Mohtashemi, K. Kleinman, and W.K. Yih, *Multi-syndrome analysis of time series using PCA: A new concept for outbreak investigation*, *Statistics in Medicine* 26 (2007), pp. 5203-5224.
- [25] Y. Okhrin, and W. Schmid, *Surveillance of Univariate and Multivariate Nonlinear Time Series*, in *Financial surveillance*, M. Frisé ed., Wiley, Chichester, 2007, pp. 153-177.
- [26] M.R. Reynolds, Jr, and K. Kim, *Multivariate Control Charts for Monitoring the Process Mean and Variability Using Sequential Sampling*, *Sequential Analysis* 26 (2007), pp. 283-315.
- [27] S.W. Roberts, *A comparison of some control chart procedures*, *Technometrics* 8 (1966), pp. 411-430.

- [28] H. Rolka et al., *Issues in applied statistics for public health bioterrorism surveillance using multiple data streams: research needs*, *Statistics in Medicine* 26 (2007), pp. 1834-1856. Available at <http://dx.doi.org/10.1002/sim.2793>
- [29] G.C. Runger, *Projections and the  $U^2$  chart for multivariate statistical process control*, *Journal of Quality Technology* 28 (1996), pp. 313-319.
- [30] G.C. Runger et al., *Optimal Monitoring of Multivariate Data for Fault Detection*", *Journal of Quality Technology* 39 (2007), pp. 159-172.
- [31] T.P. Ryan, *Statistical methods for quality improvement*, 2nd ed, Wiley, New York, 2000.
- [32] N.S. Sahni, A.H. Aastveit, and T. Naes, *In-Line Process and Product CONTROL Using Spectroscopy and Multivariate Calibration*, *Journal of Quality Technology* 37 (2005), pp. 1-20.
- [33] W.A. Shewhart, *Economic Control of Quality of Manufactured Product*, ed, MacMillan, London, 1931.
- [34] A.N. Shiryaev, *On optimum methods in quickest detection problems*, *Theory of Probability and its Applications*. 8 (1963), pp. 22-46.
- [35] C. Sonesson, *Evaluations of Some Exponentially Weighted Moving Average Methods*, *Journal of Applied Statistics* 30 (2003), pp. 1115-1133.
- [36] C. Sonesson, and M. Friséen, *Multivariate surveillance.*, in *Spatial surveillance for public health*, A. Lawson and K. Kleinman eds., Wiley, New York, 2005, pp. 169-186.
- [37] P. Wessman, *Some Principles for surveillance adopted for multivariate processes with a common change point.*, *Communications in Statistics - Theory and Methods* 27 (1998), pp. 1143-1161.
- [38] ---, *The surveillance of several processes with different change points.*, 1999:2, Department of Statistics, Gothenburg University, Göteborg, Sweden, 1999.
- [39] W.H. Woodall, and S. Amiriparian, *On the economic design of multivariate control charts*, *Communications in Statistics - Theory and Methods* 31 (2002), pp. 1665-1673.
- [40] K. Wärmefjord, *Multivariate quality control and Diagnosis of Sources of Variation in Assembled Products*, Licentiat Thesis, Göteborg University, 2004.
- [41] S. Zacks, and R.S. Kenett, *Process tracking of time series with change points*, in *Recent Advances in Statistics and Probability*, J.P. Vilaplana and M.L. Puri eds., International Science Publishers, Zeist, The Netherlands, 1994, pp. 155-171.

## Research Report

- 2007:8 Bock, D., Andersson, E. & Frisé, M.: Similarities and differences between statistical surveillance and certain decision rules in finance.
- 2007:9 Bock, D.: Evaluations of likelihood based surveillance of volatility.
- 2007:10 Bock, D. & Pettersson, K. Explorative analysis of spatial aspects on the Swedish influenza data.
- 2007:11 Frisé, M. & Andersson, E. Semiparametric surveillance of outbreaks.
- 2007:12 Frisé, M., Andersson, E. & Schiöler, L. Robust outbreak surveillance of epidemics in Sweden.
- 2007:13 Frisé, M., Andersson, E. & Pettersson, K. Semiparametric estimation of outbreak regression.
- 2007:14 Pettersson, K. Unimodal regression in the two-parameter exponential family with constant or known dispersion parameter.
- 2007:15 Pettersson, K. On curve estimation under order restrictions.
- 2008:1 Frisé, M. Introduction to financial surveillance.
- 2008:2 Jonsson, R. When does Heckman's two-step procedure for censored data work and when does it not?
- 2008:3 Andersson, E. Hotelling's T2 Method in Multivariate On-Line Surveillance. On the Delay of an Alarm.
- 2008:4 Schiöler, L. & Frisé, M. On statistical surveillance of the performance of fund managers.
- 2008:5 Schiöler, L. Explorative analysis of spatial patterns of influenza incidences in Sweden 1999 – 2008.
- 2008:6 Schiöler, L. Aspects of Surveillance of Outbreaks.
- 2008:7 Andersson, E & Frisé, M. Statistiska varningssystem för hälsorisker