



Research Report
Department of Statistics
Göteborg University
Sweden

Evaluations of some exponentially weighted moving average methods

Christian Sonesson

Research Report 2001:6
ISSN 0349-8034

| | | | |
|--------------------|-----------------------|-----------------------|---|
| Mailing address: | Fax | Phone | Home Page: |
| Dept of Statistics | Nat: 031-773 12 74 | Nat: 031-773 10 00 | http://www.stat.gu.se/stat |
| P.O. Box 660 | Int: +46 31 773 12 74 | Int: +46 31 773 10 00 | |
| SE 405 30 Göteborg | | | |
| Sweden | | | |

Evaluations of Some Exponentially Weighted Moving Average Methods

By CHRISTIAN SONESSON

Department of Statistics, Göteborg University, Sweden

SUMMARY

Several versions of the EWMA (Exponentially Weighted Moving Average) method for monitoring a process with the aim of detecting a shift in the mean are studied both for the one-sided and the two-sided case. The effects of using barriers for the one-sided alarm statistic are also studied. One important issue is the effect of different types of alarm limits. Different measures of evaluation are considered such as the expected delay, the ARL^1 , the probability of successful detection and the predictive value of an alarm to give a broad picture of the features of the methods. Results are presented both for a fixed ARL^0 and a fixed probability of a false alarm. The differences highlight the essential problem of how to define comparability between surveillance methods. The results are from a large-scale simulation study. Special attention is given to the effect on the confidence in the final results by the stochastic variation in the calibration of the methods. It appears that important differences from an inferential point of view exist between the one- and two-sided versions of the methods. It is demonstrated that the method, usually considered as a convenient approximation, is to be preferred over the exact version in many respects.

Key Words: CHANGE POINT, DETECTION, EXPECTED DELAY, EWMA, OPTIMALITY, PROBABILITY OF SUCCESSFUL DETECTION, QUALITY CONTROL, SURVEILLANCE

1 INTRODUCTION

In many areas the problem of detecting a change in a stochastic process through sequential observations is important. Examples include an increased variation in an industrial production process, an increase in the mean radiation level or an increased cancer incidence. The general goal is to detect the change in the process, occurring at an unknown time point, as quickly as possible after it has occurred and at the same time control the rate of false alarms in order to be able to take appropriate actions. Common to the situations mentioned is that a decision of whether the change has occurred in the process or not has to be made sequentially, based on the data collected so far. This means that traditional hypothesis testing cannot be used, neither can sequential tests, since we cannot stop sampling in favour of the null hypothesis. Instead we have to use statistical surveillance methods.

In this paper, the case of a positive shift in the mean of a normal distribution from one constant level, μ_0 , to another constant level, μ , at a random change point τ is considered. Let the stochastic process under surveillance be denoted by $X = \{X(t); t = 1, 2, \dots\}$ where $X(t), t \geq 1$, are assumed to be conditionally independent, given the change point, with equal variance. Without loss of generality we take $\mu_0 = 0$ and $\sigma_X = 1$. Then, $X(t) \sim N(\mu(t), 1)$, where $\mu(t) = \mu \cdot I(\tau \leq t)$ and $I(\tau \leq t)$ is the indicator function taking the value 1 if $\tau \leq t$ and 0 otherwise. At each time point, s , we want to discriminate between two states of the process, the in control state, $D(s) = \{\tau > s\}$ and the out of control state $C(s) = \{\tau \leq s\}$. To do this, we use an alarm system which consists of two parts; an alarm statistic, $p(X_s)$, where $X_s = \{X(t); t \leq s\}$, and an alarm limit, $g(s)$. The time of an alarm, t_A , is

$$t_A = \min\{s; p(X_s) > g(s)\}.$$

There are several ways to construct the alarm system. In this paper we will study different kinds of EWMA methods, all with an alarm statistic, which is an exponentially weighted moving average of the observations. The use of this alarm statistic was first introduced in statistical process control by Roberts (1959). It can be used to detect a positive shift in the mean, or a negative, or either of them. If we are only considering shifts in one direction, this is referred to as the one-sided case contrary to the two-sided case, where we are interested in shifts in either of the two directions. Several ways of constructing alarm limits to this alarm statistic has been proposed in the literature. Our focus in this paper will be on the effect of the different alarm limits on the performance of the methods.

The paper is organized in the following way. In Section 2, the different types of EWMA methods included in the study are presented. In Section 3 different ways to achieve comparability of methods are discussed. The requirements on the number of replicates in the simulation study for sufficient accuracy of conclusion is also discussed in this section. In Section 4 the comparisons of the methods for the one-sided case are presented. The one-sided case is compared with the one-sided case with a barrier and the two-sided case in Section 5. Some important differences are examined. In Section 6, some concluding remarks are given.

2 DIFFERENT TYPES OF EWMA METHODS

The EWMA method has been widely studied in the literature, mostly for various shifts in the mean in a stochastic process of normally distributed variables (Roberts (1959); Robinson and Ho (1978); Crowder (1987); Crowder (1989); Lucas and Saccucci (1990); Srivastava and Wu (1993); Chandrasekaran, English, and Disney (1995); Srivastava and Wu (1997); Schöne, Schmid, and Knoth (1999); Steiner (1999); Chan and Zhang (2000) and Frisé and Sonesson (2001)). However, the EWMA method has also been studied in many other situations. The case of a shift in the variance was studied by Chang and Gan (1994), Crowder and Hamilton (1992) and Morais and Pacheco (1998). The case of a simultaneous shift in both the mean and the variance has been studied by Domangue and Patch (1991), Gan (1995) and Morais and Pacheco (2000). Also shifts in other types of processes have been studied. Gan (1998) studied methods to detect a shift in a parameter of an exponentially distributed variable. The case of shift in the mean of an autocorrelated process was studied by Lu and Reynolds (1999), and the robustness to non-normality by Borrer, Montgomery, and Runger (1999). Multivariate EWMA methods have been studied by Tsui and Woodall (1993) and Runger and Prabhu (1996). The EWMA statistic has been shown to be useful also in other situations such as forecasting in time series (Box, Jenkins, and MacGregor (1974)).

The EWMA methods are based on an exponentially moving average, Z_s , of all accumulated observations. The alarm statistic can equivalently be represented by the recursive formula

$$Z_s = (1 - \lambda)Z_{s-1} + \lambda X(s),$$

where the weight parameter $\lambda \in (0,1]$. If $\lambda = 1$ only the last observation is used in the alarm statistic and the smaller λ is, the more equally weighted are the observations. It is well known that the standard deviation of Z_s is an increasing function of the time s and converges as s tends to infinity. The convergence will be slower for small values of λ .

$$\lim_{s \rightarrow \infty} \sigma_{Z_s} = \lim_{s \rightarrow \infty} \sqrt{\frac{\lambda}{2-\lambda} [1 - (1-\lambda)^{2s}]} = \sqrt{\frac{\lambda}{2-\lambda}}$$

What differs between the EWMA methods is the way of constructing alarm limits $g(s)$, the starting value Z_0 and whether or not a barrier for the alarm statistic is used. A barrier, b , constitutes a boundary of the alarm statistic. The one-sided version with a barrier, to detect positive shifts, uses an alarm statistic of the form $\max(b, Z_s)$. In this paper the focus is to compare the two most common ways of constructing the limits, both based on the standard deviation of the alarm statistic. The first one, here named EWMAe, uses the exact standard deviation of Z_s when constructing the alarm limits. The time of an alarm, for the one-sided case, is

$$t_A = \min\{s; Z_s > L \cdot \sigma_{Z_s}\}.$$

Since the standard deviation is increasing in time, so is the alarm limit of the EWMAe method.

The second way of constructing the alarm limits, and by far the most studied in the literature, here named EWMAa, uses the asymptotic standard deviation of Z_s and the time of an alarm, for the one-sided case, is

$$t_A = \min\{s; Z_s > L \cdot \sigma_Z\}.$$

The EWMAa method was introduced by Roberts (1959) and most of the previous studies have been concerned with this method. Robinson and Ho (1978) used an Edgeworth expansion to get a recursive technique to evaluate the average run length both in control ($ARL^0 = E[t_A | \tau = \infty]$) and out of control ($ARL^1 = E[t_A | \tau = 1]$). This was done for both the one-sided and two-sided EWMAa in discrete time. Crowder (1987) used integral equations to evaluate the properties of the run length distribution in discrete time for the two-sided EWMAa. In Crowder (1989) optimal values of the weight parameter λ to minimize the ARL^1 for a fixed ARL^0 , together with a design strategy of the method, was proposed, based on the results in Crowder (1987). Lucas and Saccucci (1990) used another approach to find the optimal value of λ for the same situation as by Crowder (1989). The alarm statistic was represented by a continuous state Markov chain. The properties were evaluated by discretizing the infinite-state transition probability matrix. For a comparison between these two approaches, see Champ and Rigdon (1991). Srivastava and Wu (1993) examined the EWMAa method in continuous time by representing it as a diffusion process. This was done for the one-sided case. Optimal values of λ to minimize the stationary average delay time, SADT, for

a fixed value of ARL^0 , were derived. After correcting for the overshoot these results were applied to the two-sided case for discrete time by Srivastava and Wu (1997), where optimal values of λ were derived for a fixed ARL^0 .

Although the types of limits used in EWMAe and EWMAa are the most common, other types have been proposed in the literature. Lucas and Saccucci (1990) suggested the use of the EWMAa method with a head start, that is $Z_0 \neq 0$. This approach assures a fast initial response to start-up problems in the process. Steiner (1999) proposed another type of fast initial response for the EWMAe method, where the usual alarm limit was multiplied by an exponential function, which lowered the alarm limit in the start. In Frisén and Sonesson (2001) the alarm limits were chosen to make the EWMA method a good approximation of a linearized version of the likelihood-ratio method for the one-sided case (see also Frisén (1999)). The likelihood ratio method gives an alarm as soon as the posterior probability of the process being out of control exceeds a fixed value. The latter two papers, Frisén (1999) and Frisén and Sonesson (2001) discussed the minimal expected delay, ED, for a fixed probability of a false alarm. The use of barriers for the EWMA alarm statistic has previously been studied in Gan (1998), where evaluations of ARL^1 were made for different values of the barrier for a fixed value of ARL^0 for the EWMAa method. This was done in the case of exponentially distributed observations.

Two-sided versions of the EWMA methods can be constructed in different ways. The most common way is to use symmetrical control limits in which case the alarm limit is constructed in the same way as for the one-sided case but instead using $|Z_s|$ as the alarm statistic. This is the approach analyzed in Section 6. Non-symmetrical alarm limits, or two parallel one-sided versions with barriers using a lower (and upper) bound of the alarm statistic, are also possible.

3 COMPARABILITY BETWEEN METHODS

3.1 *Choice of measure of false alarms*

When evaluating the effectiveness of different types of alarm systems, one has to face a trade off between false alarms and short delay times for motivated alarms. The way to handle this is usually in the same way as in a hypothesis-testing situation, where the type 1 error is fixed and evaluations of the power is made for various situations. The translation to the surveillance situation has traditionally been to characterize the type 1 error by the in control average run length to a false alarm, denoted by ARL^0 . Then different types of methods have been compared for a

fixed value of the ARL^0 . Another way of characterizing the type 1 error is by the probability of a false alarm, denoted by $P(t_A < \tau)$.

$$P(t_A < \tau) = \sum_{t=1}^{\infty} P(\tau = t)P(t_A < \tau | \tau = t)$$

The difference between these approaches is discussed by Frisén (1999) and Frisén and Sonesson (2001). In this paper evaluations of the different types of EWMA methods are carried out for both approaches. Comparisons are made for a fixed ARL^0 of 50 and 100 and also for a fixed value of $P(t_A < \tau)$ when τ is assumed to be geometrically distributed ($P(\tau = t) = v(1-v)^{t-1}$ for $t = 1, 2, \dots$) with $v = 0.05$ or 0.01 . The fixed values chosen for $P(t_A < \tau)$ corresponds to the values for the likelihood ratio method when $ARL^0 = 100$ as given by Frisén and Wessman (1999). In Section 4.1 the effects of the choice of approach will be further discussed. Results will be presented for values of λ in the interval $[0.001, 0.40]$. Small values of λ result in run length distributions with heavy right tails (for some of the situations studied here) and are therefore too computationally time consuming. Therefore only values of λ larger than or equal to 0.001 were chosen.

3.2 *Effect of uncertainty in false alarms measure*

In order to fix either the ARL^0 or the probability of a false alarm, one has to choose the constant L in the alarm limit. This can be done in several ways. In several of the earlier studies of EWMA methods the ARL^0 has been approximated as a function of L . The formulas achieved have then been used to determine the appropriate L to get a certain desired ARL^0 (Crowder (1989); Lucas and Saccucci (1990) and Srivastava and Wu (1997)). In this paper the appropriate value of L is determined by simulations. Whatever method used to determine L , careful attention to the closeness to the desired ARL^0 or the desired probability of a false alarm is needed.

The main focus in many studies, including this one, is to evaluate different methods with respect to some measure, θ . For comparability this is done under the restriction that another measure, A , has a specified value. The measure θ is usually a function of the out of control run length distribution, which depends on the constant L . The measure θ could be for example the ARL^1 for different values of λ for the EWMAa method under the restriction that A (ARL^0 or $P(t_A < \tau)$) has a specified value. Other examples of possible measures θ would be the conditional expected delay, $CED(t) = E[t_A - \tau | t_A \geq \tau = t]$, or the expected delay,

$ED = \sum_{t=1}^{\infty} P(\tau = t) \cdot CED(t)$. The aim is to study the accuracy of θ , but to do that, we start by studying the accuracy of A .

Assume that the evaluation of θ should be done for some desired value A^* of A . We aim to determine L^* such that $A(L^*) = A^*$. The procedure here to estimate L^* is to choose L_1, L_2, \dots, L_n and use simulations, with the same number m of replicates for each value of L to estimate the values of $A(L_1), A(L_2), \dots, A(L_n)$. We approximate $A(L)$ with a linear function locally, $\hat{A}(L) = \hat{a}_1 + \hat{b}_1 L$, and choose L' accordingly to give $\hat{A}(L') = A^*$. A confidence interval for L^* can be constructed by considering the test of $H_0 : A(L) = A^*$. The model for our observations is $A(L) = a_1 + b_1 L + \varepsilon$, where $\varepsilon \sim N(0, \frac{\sigma^2_{IRL}}{m})$ and σ^2_{IRL} is the variance of the in control run length, assumed to be constant locally within the range of values of L considered for the regression. First we estimate σ^2_{IRL} with such a precision that we are able to neglect the standard error of this estimate and proceed as if σ^2_{IRL} was known. For the estimation of a_1 and b_1 we use OLS. Then for the estimator $\hat{A}(L)$ we know that

$$E[\hat{A}(L)] = A(L)$$

$$V(\hat{A}(L)) = \frac{\sigma^2_{IRL}}{m} \left(\frac{1}{n} + \frac{(L - \bar{L})^2}{\sum_{i=1}^n (L_i - \bar{L})^2} \right).$$

To test the hypothesis $H_0 : A(L) = A^*$ we use the test statistic $K(L) = \frac{\hat{A}(L) - A^*}{\sqrt{V(\hat{A}(L))}} \sim N(0,1)$,

which can also be used for the construction of confidence intervals for L^* . We denote the p :th percentile in the $N(0,1)$ distribution by z_p . Then, included in a two-sided confidence interval $[L_l, L_u]$ for L^* of confidence level $(1-\alpha)$ will be those L for which $K^2(L) \leq z^2_{1-\alpha/2}$ where we have

$$L_l = L' - \frac{z_{1-\alpha/2} \sigma_{IRL}}{\hat{b}_1 m} \sqrt{\frac{1}{n} + \frac{(L_l - \bar{L})^2}{\sum_{i=1}^n (L_i - \bar{L})^2}}$$

$$L_u = L' + \frac{z_{1-\alpha/2} \sigma_{IRL}}{\hat{b}_1 m} \sqrt{\frac{1}{n} + \frac{(L_u - \bar{L})^2}{\sum_{i=1}^n (L_i - \bar{L})^2}}.$$

From the confidence interval $[L_l, L_u]$ for L^* we can also construct a confidence interval for $L'-L^*$ as

$$\left[-\frac{z_{1-\alpha/2} \sigma_{IRL}}{\hat{b}_1} \frac{\sigma_{IRL}}{m} \sqrt{\frac{1}{n} + \frac{(L_l - \bar{L})^2}{\sum_{i=1}^n (L_i - \bar{L})^2}}, \frac{z_{1-\alpha/2} \sigma_{IRL}}{\hat{b}_1} \frac{\sigma_{IRL}}{m} \sqrt{\frac{1}{n} + \frac{(L_u - \bar{L})^2}{\sum_{i=1}^n (L_i - \bar{L})^2}} \right].$$

What we are interested in is $\theta(L^*)$ and our aim is to construct confidence intervals for the difference $\theta(L') - \theta(L^*)$. However, we do only have an estimate, L' , of L^* . After choosing a constant L' , we have an estimate $\hat{\theta}(L')$ of $\theta(L')$, and knowledge about its stochastic properties for this fixed value of L .

To construct a confidence interval for $\theta(L') - \theta(L^*)$ we approximate $\theta(L)$ by a linear function, $\theta(L) = a_2 + b_2 L$ locally around L' . Simulations of $\theta(L)$ using values $L_1, L_2, \dots, L_{n'}$ of L around L' can achieve confidence intervals for b_2 . For each value of L_i , $i = 1, \dots, n'$, let m'_i be the number of replicates that the estimate $\hat{\theta}(L_i)$ is based on. (For several possible θ , for example the $CED(t)$, the evaluation is based on the alarms for which $t_A \geq \tau$. The reason for the number of replicates to be unequal is due to the difficulty of dimensioning the simulations to achieve the same number of alarms at or after the change point for each value of L . Note though that for $\tau = 1$ this problem does not arise.) The regression parameters a_2 and b_2 will be estimated using WLS since the number of replicates for each value of L_i , $i = 1, \dots, n'$ is not the same. Our observations follow the model $\theta(L_i) = a_2 + b_2 L_i + \delta_i$, where $\delta_i \sim N(0, \frac{\sigma^2_{ORL}}{m'_i})$ and σ^2_{ORL} is the variance of the out of control run length (only considering the alarms for which $t_A \geq \tau$). Also in this case we regard σ^2_{ORL} as known after estimating it with high precision. The WLS estimate of b_2 has the properties

$$E[\hat{b}_2] = b_2$$

$$V(\hat{b}_2) = \sigma^2_{ORL} \frac{n'}{n' \sum_{i=1}^{n'} m'_i L_i^2 - (\sum_{i=1}^{n'} \sqrt{m'_i} L_i)^2}.$$

We can then construct confidence intervals for b_2 using that $\frac{\hat{b}_2 - b_2}{\sqrt{V(\hat{b}_2)}} \sim N(0,1)$.

Let the constructed confidence interval, Λ , for $(L' - L^*)$ be of confidence $(1 - \alpha_1)$ and the constructed confidence interval, B , for b_2 be of confidence $(1 - \alpha_2)$. Then we can combine these intervals to construct a confidence interval of confidence at least $(1 - \alpha_1)(1 - \alpha_2)$ for $\theta(L') - \theta(L^*) = b_2(L' - L^*)$ taking $\min\{b \cdot l, b \in B, l \in \Lambda\}$ to be the lower limit and $\max\{b \cdot l, b \in B, l \in \Lambda\}$ to be the upper.

Confidence intervals for $\theta(L') - \theta(L^*)$ constructed in this way for some cases studied in the paper can be found in Table 1. Included are those situations where the variances in the in control run length distributions are the largest. These confidence intervals indicate that the determination of L' is good enough to guarantee reliable comparisons between the methods. The conclusion is that the numbers of replicates in the simulations are enough (but not unnecessary) for the present purposes.

Table 1. Confidence intervals for $CED(L') - CED(L^*)$ when $\mu = 1$, $ARL^0 = 50$, $n \geq 10$, $n' = 10$, $m = 500000$. Level of confidence = 0.9025.

| Method | $\tau = 1$ ($m' = 1000000$) | | $\tau = 20$ ($m' \geq 143500$) | |
|-----------------------|----------------------------------|-------------|-------------------------------------|-------------|
| | Lower limit | Upper limit | Lower limit | Upper limit |
| EWMAa(0.001), 1 sided | -0.0008 | 0.0012 | -0.0008 | 0.0013 |
| EWMAe(0.001), 1 sided | -0.0008 | 0.0008 | -0.0055 | 0.0056 |
| EWMAa(0.001), 2 sided | -0.0009 | 0.0029 | -0.0009 | 0.0029 |
| EWMAe(0.001), 2 sided | -0.0023 | 0.0023 | -0.0044 | 0.0045 |

4 COMPARISONS FOR THE ONE-SIDED CASE

We start by the comparisons between the EWMAa and the EWMAe methods for the one-sided case. In Section 5, the focus will be on the difference between constructing the methods for the one-sided case, the one-sided case with a barrier and two-sided case. Our main interest is to see the effect of the different alarm limits on the performance of the methods, both when the process is in- and out of control.

4.1 *On the difference between a fixed ARL^0 and a fixed probability of a false alarm*

The only thing that differs between the EWMAa and the EWMAe methods is the alarm limits. As already mentioned, the variance of the alarm statistic is an increasing function of time and converges as the time tends to infinity. This means that in order to fix the in control run length or the probability of a false alarm, the value of L in the alarm limit will be larger for the EWMAe than for EWMAa method, for a fixed value of λ . Since the convergence is slower for small values of λ the difference in L will be most pronounced for small values of λ . The difference in the alarm limits between the methods will effect their relative performances at different time points, both with respect to their false alarm distributions as well as the ability to detect the change at different time points as will be shown later.

An important thing is to notice the difference between using a fixed ARL^0 compared with a fixed probability of a false alarm. Since the in control run length distributions are different, an equal ARL^0 does not imply equal probability of a false alarm and vice versa. In Figure 1 the values of ARL^0 for a fixed value of the false alarm probability are presented.

For moderate values of λ , the value of ARL^0 is fairly constant if we have fixed the probability of a false alarm for both methods. There is only a slight difference between the EWMAa and EWMAe methods for these values of λ . However the required ARL^0 to fix the probability of a false alarm is much larger for smaller values of λ . This means a larger L for the case of a fixed probability of a false alarm for both methods. For small values of λ there is a large difference between the methods. We will therefore expect both the in control as well as the out of control performance to differ depending on the choice of a fixed ARL^0 or a fixed probability of a false alarm. We also expect different results for different values of λ . The difference in performance depends on our choice of ν . This can be seen comparing Figure 1a and 1b.

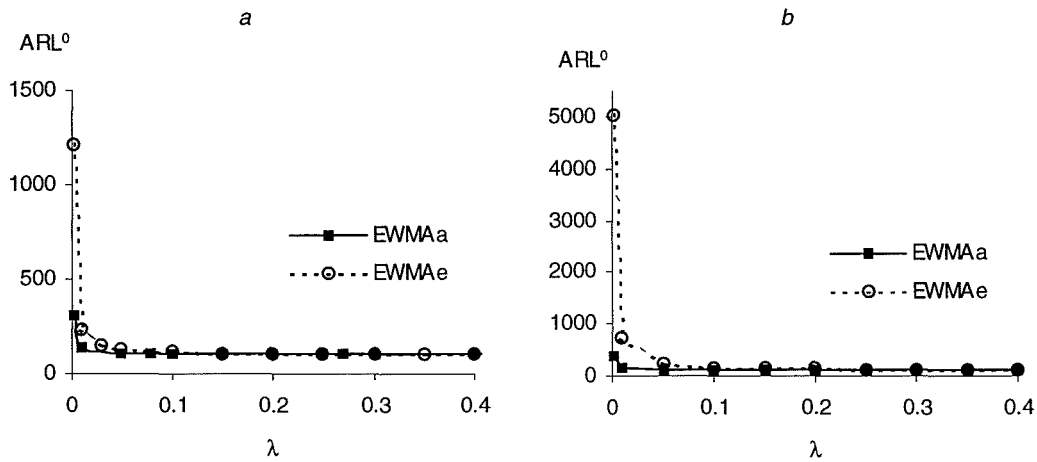


FIG 1: ARL^0 as a function of λ .
a. $P(t_A < \tau) = 0.4877$ for $v = 0.01$., b $P(t_A < \tau) = 0.1326$ for $v = 0.05$.

The difference in ARL^0 for a fixed $P(t_A < \tau)$ reflects the importance of defining what is meant by comparability between methods of surveillance. The choice of fixing ARL^0 or $P(t_A < \tau)$ and at what level will affect the constants in the alarm limits. For the situations of a fixed $P(t_A < \tau)$ studied here, and shown in Figure 1, the adjustments needed in the value of L (measured as the increase in the ARL^0) when choosing a fixed $P(t_A < \tau)$ compared with a fixed value of $ARL^0 = 100$, is larger for EWMAe than for EWMAa and larger for small than for large values of λ . The values of $P(t_A < \tau)$, for which the methods have been fixed, are the ones, which result in a value of 100 for the ARL^0 for the likelihood ratio method (Frisén and Wessman (1999)), which can be used as a benchmark. From Figure 1, we can expect that the EWMAa method will perform relatively better than EWMAe and large values relatively better than small values of λ if we chose to fix $P(t_A < \tau)$ instead of ARL^0 . If this pattern is true in general with other values of $P(t_A < \tau)$, or another distribution of τ , is not examined here. The reason for expressing the difference between the cases in terms of the increase in ARL^0 in the case of a fixed $P(t_A < \tau)$ here is the fact that ARL^0 is the conventional measure of in control behavior and results expressed in this form can be easily interpreted. We can also consider how the alarm probabilities at different time points are affected. This comparison can be found in Figure 2 and 3.

In some cases there has been argued that ARL^0 is the only necessary in control characteristic with the motivation that the in control run length distribution is approximately geometric. The results here indicate that this is not the case, as will be seen in the figures displaying the in control run length distributions. What is clear though is that the difference between the methods in both in

control features as well as detection abilities will depend on our choice of in control characteristic. The most common way is to fix the ARL^0 . However, one should be aware of the consequences of such a choice.

4.2 In control features for the one-sided case

No matter whether the ARL^0 or the probability of a false alarm is fixed, the alarm limit for EWMAe will be lower than the one for EWMAa for early, but larger for later time points, for a fixed value of λ , due to the difference in time dependency of the alarm limits. The time point at which the alarm limits cross is later for small values of λ reflecting the slow convergence of the variance.

In Figure 2 we can see the effect of the different alarm limits on the in control run length distributions when $ARL^0=100$. The EWMAe method has a higher probability of a false alarm than the EWMAa method for a fixed value of λ at early time points since the alarm limits are lower. We can also notice a deviation from a geometric distribution, especially for the EWMAa method.

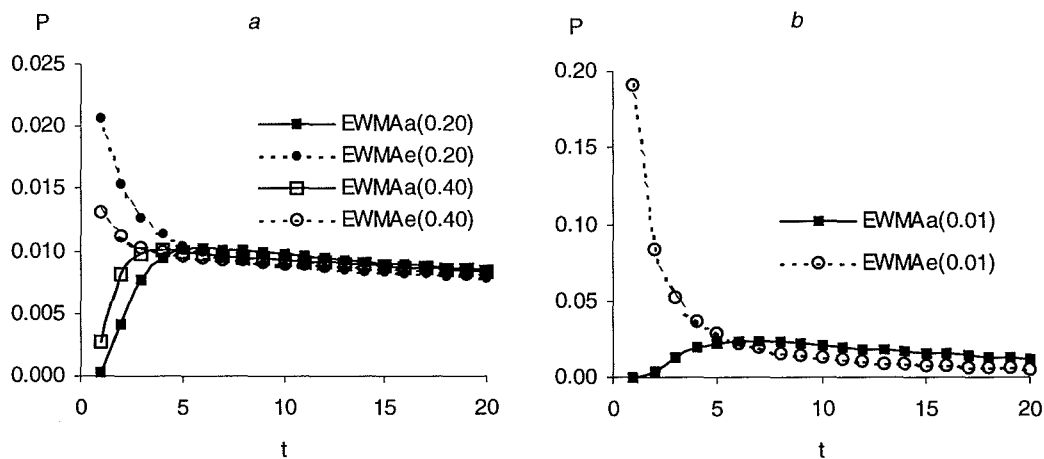


FIG 2: In control run length distribution when $ARL^0 = 100$. In the figure, the value of λ is indicated in the parentheses.

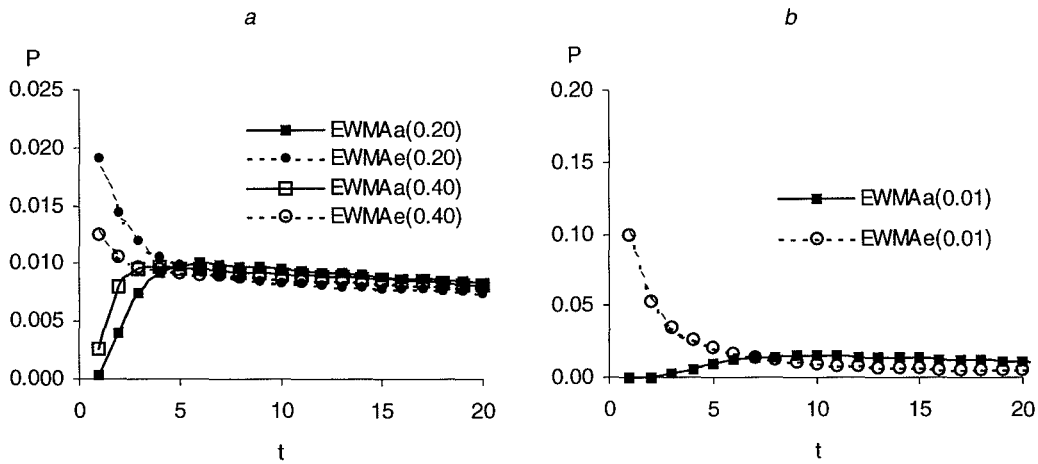


FIG 3: In control run length distribution when $P(t_A < \tau) = 0.4877$ for $\nu = 0.01$.

The in control run length distribution, when instead the probability of a false alarm is fixed, is illustrated in Figure 3 for the case when $\nu = 0.01$. We note especially the decreased probability of a false alarm at the first time point, comparing Figure 2b with Figure 3b, for the EWMAe when $\lambda = 0.01$ as a result of an increased value of L .

4.3 Detection of a true change for the one-sided case

There are several ways suggested in the literature to compare the performances of different surveillance methods with respect to true changes in the process. In the area of statistical process control the focus has traditionally been on the ARL^1 with optimality defined as minimal ARL^1 for a fixed value of ARL^0 . This criterion has also been the common one when designing the EWMA methods (Crowder (1989); Lucas and Saccucci (1990) and Srivastava and Wu (1997)). However, as an optimality criterion, this is not without critics (Frisén (1999) and Frisé and Sonesson (2001)). Gan (1993) considered instead the median out of control run length due to the skewed run length distributions. However, also in this case only shifts at the first time point were considered. This approach might be reasonable in an industrial manufacturing process, where one suspects various start-up problems. However, in the overwhelming majority of applications the possibility of later shifts should also be taken into account. One example of this is the monitoring of the foetal heart rate during labour (Frisén (1992)), where the foetus can suffer from a lack of oxygen. This can happen at any time point during the labour, which normally takes many hours, and thus in this case we must also take into account possible late shifts. Other examples include the surveillance of radiation levels (Järpe (2000)) and the surveillance of diseases (Sonesson and Bock (2001)). Since the false alarm probabilities at different time points differs between the methods, so will the

detection ability of a true change. Only considering shifts at the first time point will favour EWMAe over EWMAa with respect to the detection ability. This is because the false alarm distributions are different, as could be seen in Figure 2 and 3, where EWMAe gives more false alarms at early time points than EWMAa for the same value of λ .

In this paper we will use several kinds of measurements in order to get a broad picture of the ability of the methods to detect a true shift. Here we will consider the ARL^1 , the conditional expected delay, the expected delay, the probability of successful detection and the predictive value of an alarm.

4.3.1 A change at the first time-point for the one-sided case

As expected, due to the difference in the false alarm distributions, the EWMAe method has shorter ARL^1 compared with the EWMAa method for the same value of λ as can be seen in Figure 4, where $\mu = 1$ for $ARL^0=50$ and 100. Results for $\mu = 0.5$ can be found in Section 5.

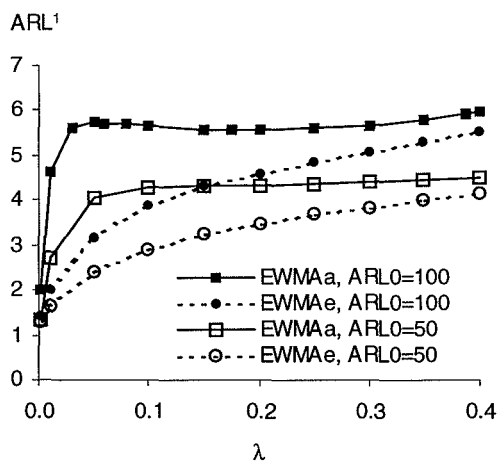


FIG 4: ARL^1 as a function of λ for $\mu = 1$ when $ARL^0=50$ and $ARL^0=100$.

Figure 4 indicates that in order to minimize the ARL^1 for a fixed value of the ARL^0 λ should approach zero (although the studied values of λ only covered the interval $[0.001, 0.40]$) both for EWMAa and EWMAe independently of the value of ARL^0 . Then, old and new observations have the same weight in the alarm statistic, and thus one of the necessary conditions given in Frisén (1999) for minimizing the ARL^1 is fulfilled. This will further be elaborated in Section 5, when comparing the results for the one-sided case with the two-sided case.

4.3.2 *Detection of later changes for the one-sided case*

The ability to detect changes at later time points, t , can be evaluated with respect to the conditional expected delay, $CED(t) = E[t_A - \tau | t_A \geq \tau = t]$. In the case of a change at the first time point, $CED(1) = ARL^1 - 1$. By considering the conditional expected delay, we are no longer limited to changes occurring at the first time point.

In Figure 5, the characteristics of the CED as a function of the time of the change can be examined. The EWMAe method has low delay times if the change occurs in the first time points. However the delay time increases with the time of the change. The results in Figure 5 clearly indicates that it is not enough only to consider changes at the first time point when evaluating EWMA methods since the time point of the change plays a crucial role in the detection ability. The increase in the observed values of CED is the largest for small values of λ . This holds both for the EWMAa and the EWMAe method and is the price to pay for the low delay times for early changes.

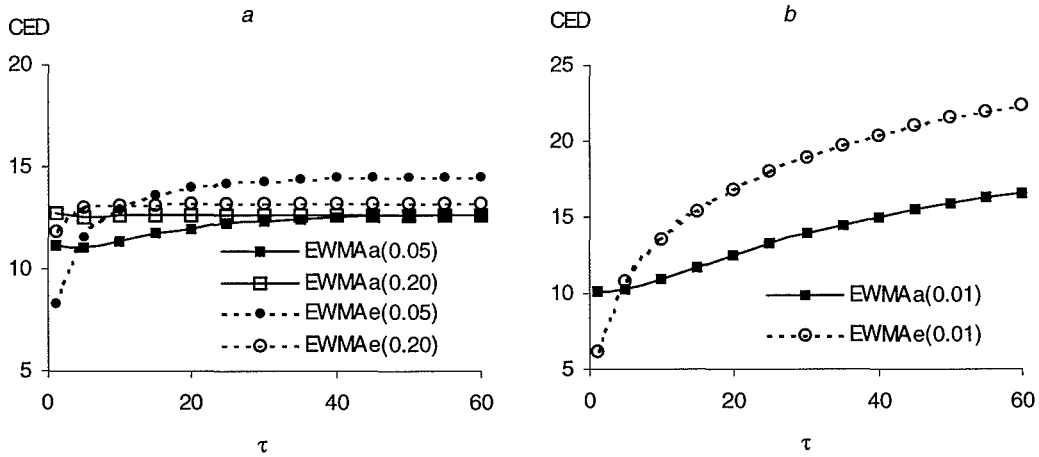


FIG 5: CED as a function of τ for $\mu = 0.5$ when $P(t_A < \tau) = 0.4877$ for $\nu = 0.01$.

To summarize the CED values for different time points different approaches can be taken. Two ways will be described here. The first one focuses on minimax properties of the methods with respect to the CED for various time points. The other way is to average the CED with respect to the distribution of the change-point τ .

Minimax criteria can be defined in various ways. Focus is on the maximum value of the conditional expected delay, which should be minimized. This type of evaluation has been studied

extensively in the literature (Lorden (1971), Pol lak (1985), Moustakides (1986)). One surveillance method known to possess optimality characteristics when the minimax criterion is expressed as minimal maximal conditional expected delay with respect to τ and the worst possible outcome of $X_{\tau-1}$ for a fixed value of the ARL^0 is the CUSUM method (Moustakides (1986)). In general, the CUSUM method is defined by the stopping rule

$$t_A = \min(s; w_s - \min_{0 \leq j \leq s} w_j > K),$$

where $w_s = \sum_{t=1}^s y_t$, $y_t = \log(f_1(X(t)) / f_0(X(t)))$ and f_0 and f_1 denotes the in control and out of control distributions.

For the case studied in this paper, the time of an alarm for the CUSUM method can also be written recursively as

$$t_A = \min\{s; S_s > h\},$$

where $S_t = \max(0, S_{t-1} + X(t) - \mu/2)$ and $S_0 = 0$. Thus, the alarm statistic for the CUSUM method has a lower boundary of zero. For EWMA based methods, not much attention has been drawn to minimax evaluations. In Gan (1995) and Gan (1998) this has been discussed for the detection of different types of changes in exponentially distributed variables.

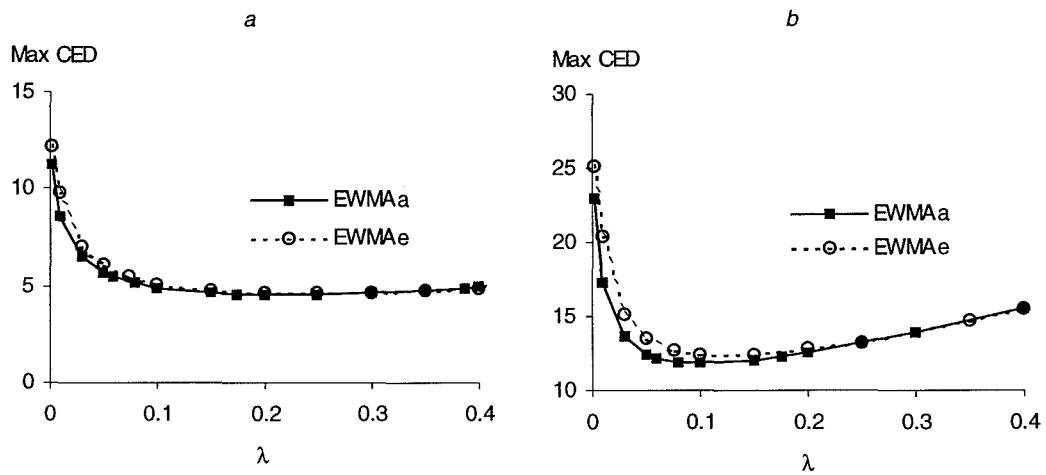


FIG 6: Maximal CED as a function of λ when $ARL^0=100$.

a. $\mu = 1$, b. $\mu = 0.5$.

In Figure 6 we see the effect of the different alarm limits on the maximum value of CED (with respect to the time point of the change) for a fixed value of the $ARL^0=100$. In the case when $\lambda \leq 0.01$ (where $CED(t)$ is still increasing with t when $t=100$) the values of $CED(100)$ are presented. In this sense, the EWMAa method is superior to the EWMAe. However, the difference between the methods is smaller if μ is large.

When summarizing the conditionally expected delay at different time points with respect to the distribution of the change-point τ , the standard procedure is to assume that τ is geometrically distributed with parameter ν . This implies a constant intensity of a shift, $\nu = P(\tau = t | \tau > t - 1)$. However, this assumption might be questioned in many applications. Assuming a distribution for τ , the expected delay is defined as

$$ED = E_{\tau} [t_A - \tau | t_A \geq \tau] = \sum_{t=1}^{\infty} P(\tau = t) \cdot CED(t).$$

The use of minimal expected delay for a fixed probability of a false alarm has been suggested as an optimality criteria and leads to the likelihood ratio method (Frisén and de Maré (1991)). This method is equivalent of making an alarm as soon as the posterior probability of an alarm exceeds a fixed value. The assumption made with respect to ν will determine the method and parameters which minimize the expected delay in the class of EWMA methods. If ν is large, methods with low $CED(t)$ for early time points will be preferable. If on the other hand ν is close to zero, methods with low values of $CED(t)$ at late time points will be preferred. Specifically, the value of ν will determine which of the EWMAa or EWMAe alarm limits that will be preferable as well as what value of λ that will minimize the expected delay. The expected delay can also be considered if the ARL^0 is fixed as is done in Figure 7.

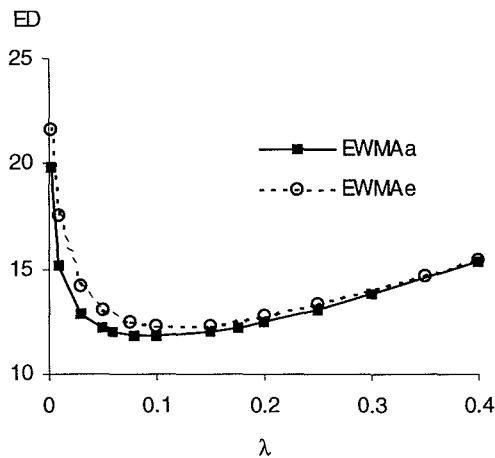


FIG 7: $ED(\nu = 0.01)$ as a function of λ for $\mu = 0.5$ when $ARL^0 = 100$.

For the case illustrated in Figure 7, the EWMAa method has lower expected delay than the EWMAe method. However, if ν is large enough, the EWMAe method has lower expected delay for all cases studied here. Consider for example the case when $\nu = 1$, in which case only a shift at the first time point will be of interest and $ED = ARL^1 - 1$ and then EWMAe will have a lower expected delay as could be seen in Figure 4. However, a value of $\nu = 0.01$ is not large enough for the EWMAe method to have a lower expected delay than the EWMAa method. In the majority of applications we expect that the EWMAa method will be preferable to the EWMAe method with respect to the expected delay. The suggested versions of the alarm limits to insure a fast initial response to start-up problems (Lucas and Saccucci (1990) and Steiner (1999)) can be expected to have even larger expected delay than the EWMAe method for cases with small values of ν . The difference between the EWMAa and the EWMAe method is larger, in absolute value, for a smaller change (not illustrated). This is in accordance with results in Frisén and Wessman (1999) (in that case regarding likelihood-based methods) that methods are alike if the change is large.

If instead we use a fixed value of the probability of a false alarm, the difference between the EWMAa and the EWMAe method is larger than for the case of a fixed ARL^0 , which is presented in Figure 8.

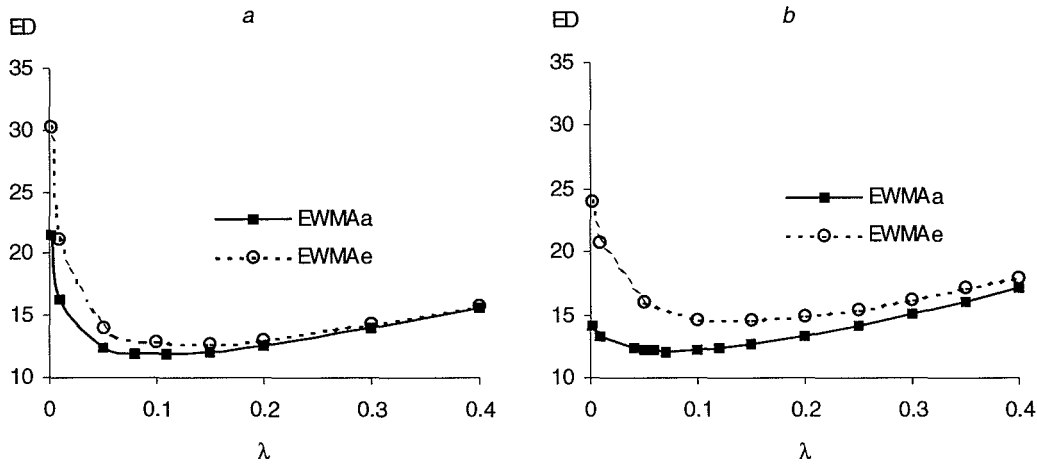


FIG 8:

a. $ED(v=0.01)$ as a function of λ for $\mu=0.5$ when $P(t_A < \tau) = 0.4877$ for $v=0.01$.

b. $ED(v=0.05)$ as a function of λ for $\mu=0.5$ when $P(t_A < \tau) = 0.1326$ for $v=0.05$.

Common to the measures of performance for detection of a true shift considered so far (ARL^1 , CED and ED) is that they focus on the mean of the out of control run length distribution conditioned on the time point of the shift. However, as in the in control case discussed previously, the mean does not account for all information in the out of control run length distribution. One alternative measure is the median delay time, considered in Gan (1993) for a shift at the first time point.

In some cases, the application considered calls for other types of evaluations than the mean delay time, for example the case where a limited time for actions exists. An example is the outbreak of an infectious disease where an epidemic starts if no actions are taken, or the case of surveillance of a foetus heart rate. For those cases, the expected value of the delay is of less interest. Instead the probability of successful detection, PSD, defined as the probability of detecting a change within a certain time interval, d , after a true change, is more important (Frisén (1992)).

$$PSD(d, t) = P(t_A - \tau \leq d \mid t_A \geq \tau = t)$$

The PSD is thus a function both of the time of the change and the length of the interval in which the detection is defined as successful. In Figure 9, the PSD is presented as a function of d for a fixed value of τ for the case of a fixed probability of a false alarm when $v=0.01$.

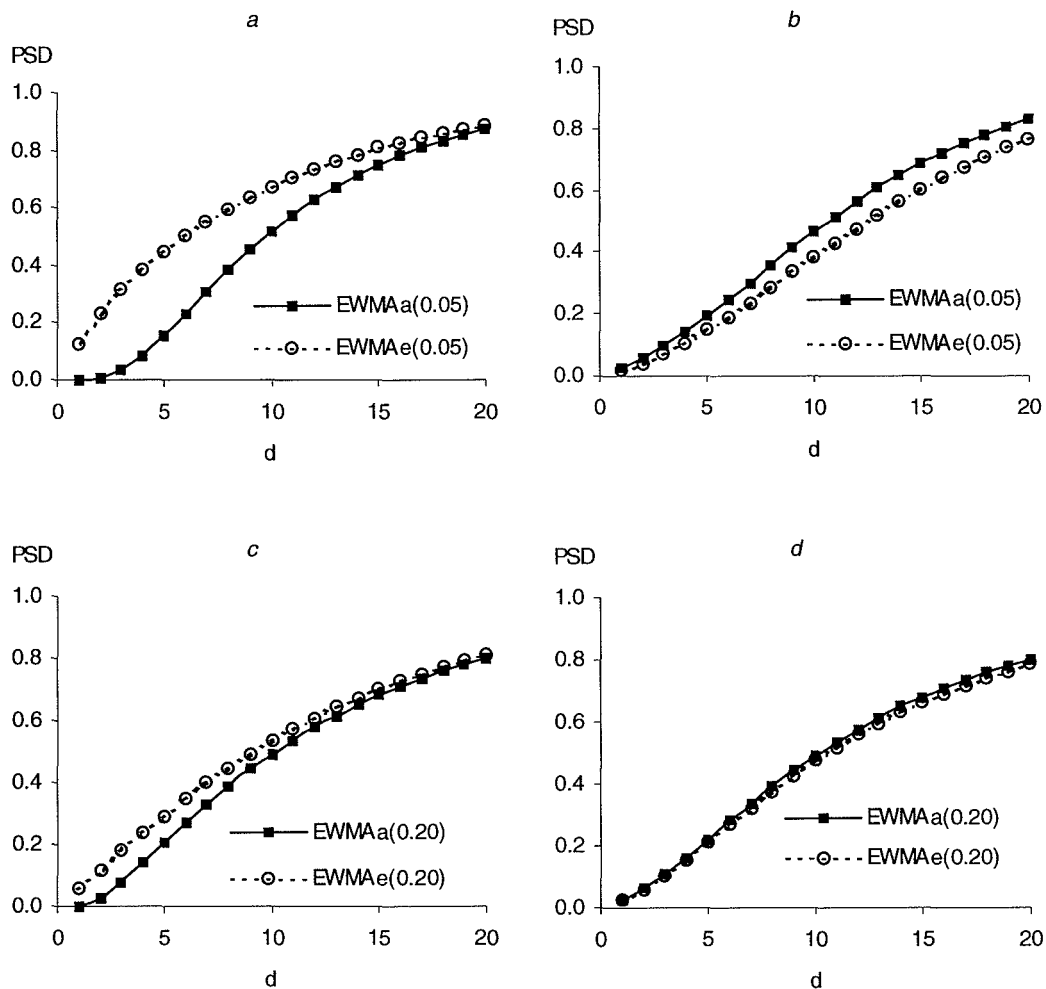


FIG 9:

PSD as a function of d for a fixed value of τ for $\mu = 0.5$ when $P(t_A < \tau) = 0.4877$ for $\nu = 0.01$.

a. $\tau = 1$, b. $\tau = 20$,

c. $\tau = 1$, d. $\tau = 20$.

For the case in Figure 9, the EWMAa is preferable to the EWMAe if the change occurs at a later time point. However, if the change occurs at the first time point the relationship is reversed.

4.4 The confidence to put in an alarm in the one-sided case

One factor, which is often neglected in the evaluation of surveillance methods, is most important when applying the method, namely what to do if an alarm is triggered. The answer to this question depends of course on the application considered. However, when constructing a

method of surveillance, this should be kept in mind. Of major importance is what degree of belief to put in an alarm. In Frisén (1992), the predictive value, $PV(t) = P(C(t) | t_A = t)$ of an alarm was suggested as an evaluation criterion. The motivation is that an alarm with low predictive value should not cause the same actions as one with high predictive value. For the coordination of the actions to follow an alarm, it is preferable if the predictive value is approximately constant over time.

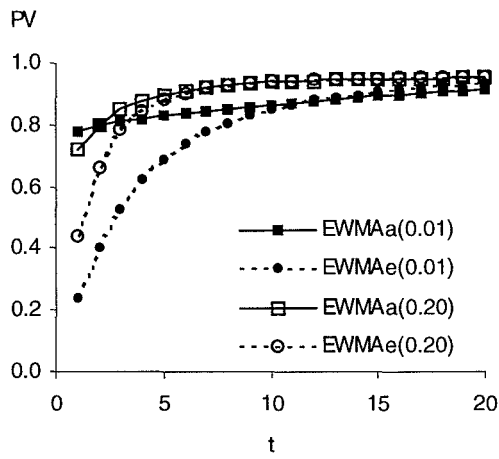


FIG 10:

PV as a function of time for $\nu = 0.10$, $\mu = 1$, when $ARL^0 = 100$.

One effect of the short delay times for early changes for the EWMAe method is a low predictive value of an alarm at early time points, especially for small values of λ , as illustrated in Figure 10. The EWMAa method has more attractive predictive value features for this case. When $\lambda = 0.20$, the predictive value is not far from constant.

5 TWO-SIDED CASE AND BARRIERS

In this section we will first focus on the different optimal values of λ to minimize the ARL^1 for a fixed value of ARL^0 between the one- and two-sided versions. Secondly, we will consider the minimax properties and also include the one-sided version with a barrier for comparison.

The two-sided version is used in order to detect both positive and negative shifts. Here we study the case of symmetrical alarm limits around 0 using $|Z_s|$ as the alarm statistic. The most obvious difference from the one-sided case is that the value of the alarm statistic at each time point

(conditioned on no alarm) is bounded downwards by the lower alarm limit. This is also the case when using a barrier for the one-sided version. In the usual one-sided version no lower boundary exists. The one-sided version with a barrier to detect positive shifts uses an alarm statistic of the form $\max(b, Z_s)$, where b constitutes the barrier or the lower bound of the alarm statistic. Barriers have important consequences, as will be explored below.

In simulations, we will consider fixed values of ARL^0 of 50 and 100. The constant, L , in the alarm limit will be larger when using a barrier or two-sided alarm limits than for the one-sided version for the same value of λ . However, the difference depends on the value of λ and also on the kind of alarm limits used. The extent to which L is altered will impose a difference in the appearance between the one-sided, the one-sided with a barrier and the two-sided versions of the methods. Worth noting is the agreement of the values of L in this study with those given in Crowder (1989) for the two-sided version of the EWMAa method and thus, these simulations confirm the values given there.

5.1 Differences in the in control features

In Figure 11, the in control run length distributions are presented both for the one- and two-sided case when $ARL^0=50$.

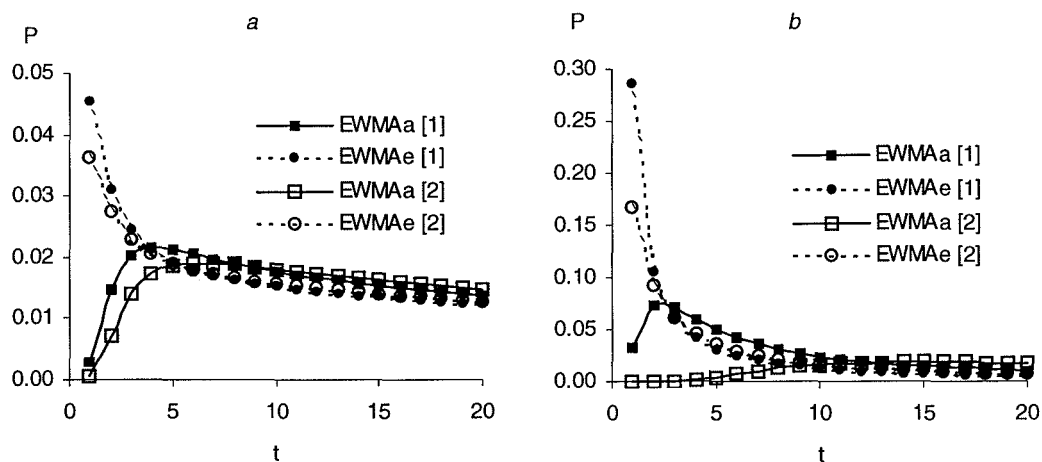


FIG 11:

In control run length distribution when $ARL^0=50$. The value in [] indicates if it is the one- or two-sided version considered.

a. $\lambda = 0.20$, b. $\lambda = 0.01$.

The common feature is that the probability of an early false alarm for the two-sided case is lower than for the corresponding one-sided case. The error spending has thus been shifted towards later time points, due to a change in the dependency structure between successive decisions.

A difference between the EWMAa and the EWMAe method is the mode of the in control run length distribution. For the EWMAe method, the mode is 1 for both the one- and two-sided case. For the EWMAa method, the mode is larger for the two-sided case, especially for small values of λ . This indicates a change also in the ability to detect a true change for the EWMAa method with small values of λ .

5.2 *The differences in detecting a true change*

We now focus on the detection of a true change and explore the differences between the one-sided, the one-sided with a barrier and the two-sided versions of the EWMA methods.

5.2.1 *A change at the first time-point*

When we are in a surveillance situation only considering the ARL^1 , all we have to decide is whether all observations are from the in control or out of control distribution. For the one-sided case, with fixed value of the ARL^0 , this implied equal weight to all observations in order to minimize the ARL^1 . We can compare this situation with a hypothesis test, using a fixed sample, and a sequentially hypothesis testing situation. Both for the hypothesis test and the sequential hypothesis test we want to decide which of the two possible distributions all sampled observations come from.

In a one-sided hypothesis test situation with a fixed sample, equal weight will also be given to all observations for the optimal method. In that case, optimality is usually defined as maximal power for a fixed significance level, as in the Neyman-Pearson Lemma. For a one-sided test with specified means and known variances (simple null and alternative hypotheses) the resulting test statistic is the mean of the observations. This is also the test statistic of the uniformly most powerful test in the case of a composite alternative hypothesis $\mu > \mu_0$. For the two-sided case though, the situation is not that simple. No uniformly most powerful test exists.

In a sequential test situation optimality is often defined in terms of minimal expected sample size, both under H_0 and H_1 , among all tests having no larger error probabilities. In the case of simple null and alternative hypothesis, the resulting test is the SPRT, sequential probability ratio

test. In the case of two normal distributions with specified mean and known variance, the SPRT also results in an optimal test statistic that is the mean of the sampled observations. In this case no uniformly most powerful test exists in the two-sided case.

Therefore, the result in the one-sided case that the value of λ , that minimizes the expected number of sampled units, gives equal weight to all observations in the alarm statistic, independent of the ARL^0 for both the EWMAa and the EWMAe method, should not be surprising.

Many of the previous studies of EWMA methods have been determined to minimize the ARL^1 for a fixed value of ARL^0 for the two-sided case. In Figure 12, values of the ARL^1 are given for a fixed value of ARL^0 . The simulations support the results by numerical approximations in Lucas and Saccucci (1990), Crowder (1989) and Srivastava and Wu (1997) concerning the two-sided EWMAa method with respect to the optimal weighting, λ , of the observations. However, the values presented here suggest that the approximations used in the previous papers overestimate the value of the optimal ARL^1 slightly.

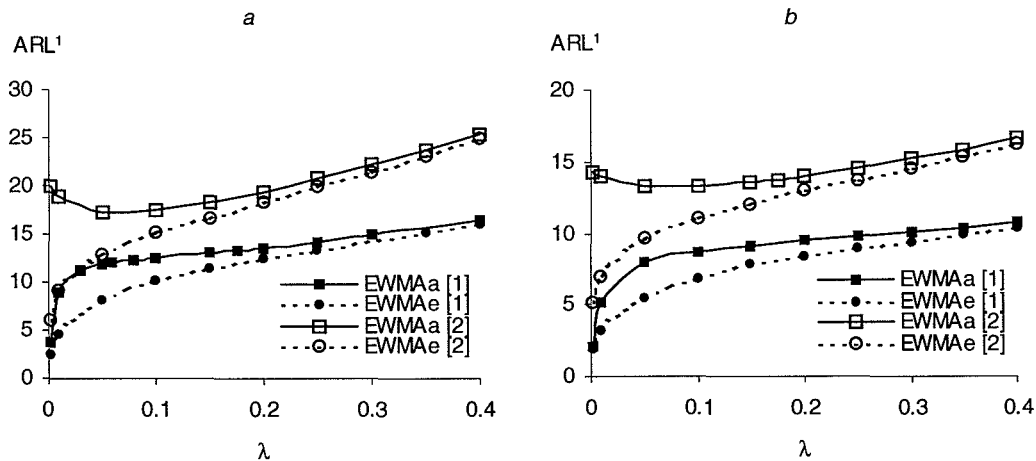


FIG 12: ARL^1 as a function of λ for $\mu = 0.5$.

a. $ARL^0 = 100$. b. $ARL^0 = 50$.

An interesting feature is that the optimal value of λ differs considerably between the EWMAa and EWMAe method for the two-sided case, which was not the case in the one-sided situation. For the EWMAe method the optimal λ still implies equal weight to all observations in the two-sided case. This is no longer the case for the EWMAa method. The difference between the one-sided and two-sided EWMAa is the result of the different error spending as a result of the lower boundary that the two-sided alarm limits imply, which could be seen also in Figure 11 of the in control run

length distributions. However, using a value of λ not equal to zero to distinguish between the in control and out of control distribution, when all observations come from either of them, violates fundamental inference principles.

5.2.2 Detection of changes at later time-points

The use of the lower bound for the alarm statistic will also affect the conditional expected delay of the methods. Here, we focus on the small values of λ and in Figure 13 results are given for $\lambda = 0.01$. For the one-sided case, CED is increasing with time both for the EWMAa and the EWMAe method. For the two-sided case, CED is approximately constant but slightly decreasing over time for the EWMAa while for the EWMAe, CED is increasing with time also for the two-sided case. For both methods, the dependency on time for CED is similar for the two-sided version and the one-sided version with a barrier. Both these versions have a bounded alarm statistic.

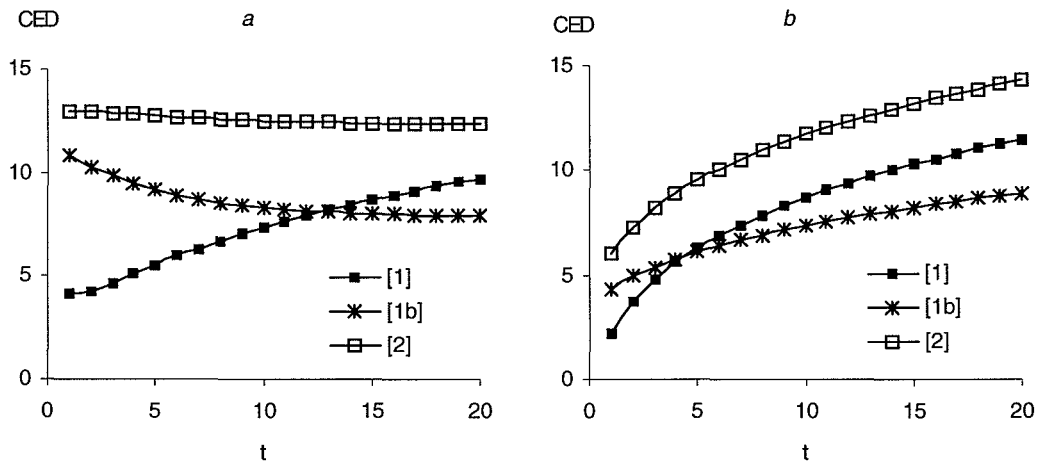


FIG 13: CED as a function of time for $\mu = 0.5$ when $ARL^0 = 50$. The barrier, b , is set to zero in the figure.

a. EWMAa(0.01), b. EWMAe(0.01).

The lower boundary of the alarm statistic for the one-sided case with a barrier and for the two-sided case will also affect the maximum value of CED of the methods. Figure 14a illustrates the maximum value of the CED for the one-sided and two-sided cases when $ARL^0 = 50$ and $\mu = 0.5$. In the case when $\lambda \leq 0.01$ for the one-sided versions, as well as for the two-sided EWMAe (where $CED(t)$ is still increasing with t when $t = 100$), the values of $CED(100)$ are presented. An interesting difference can be seen between the EWMAe and the EWMAa method. For the EWMAa method the two-sided version has considerably lower maximal CED for small values of

λ than the one-sided version. For the EWMAe method this is not true. For the two-sided EWMAa method, the maximal CED is attained at the first time point for all values of λ (see also Figure 13a). In Figure 14b the same situation is illustrated for the EWMAa method with different values of the barrier. Note that the one-sided version is equal to a barrier at $-\infty$. As indicated in Figure 14b, different values of the barrier will be preferred for different values of λ .

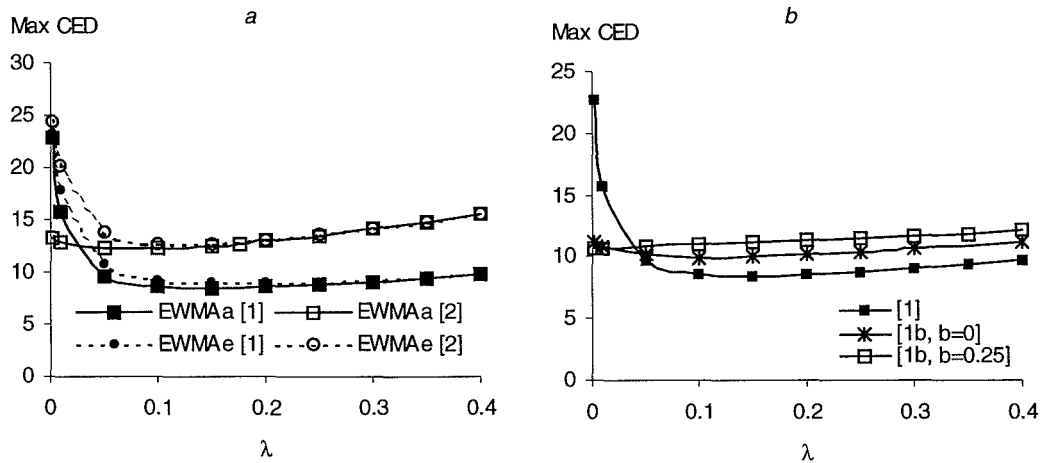


FIG 14: Maximal CED as a function of λ for $\mu = 0.5$ when $ARL^0 = 50$.

6 CONCLUDING REMARKS

The surveillance of a random process to detect a shift in the process has wide-spread application possibilities. To be able to make correct decisions about the state of the process at each time point the help of a properly designed surveillance system is needed. In this paper we have studied different EWMA methods. The focus has been on the effect of different types of alarm limits.

To the EWMA statistic various forms of alarm limits have been suggested in the literature. The most common ones are the EWMAa and the EWMAe. However, the comparative studies between these have only considered shifts at the first time point. In that case, the EWMAe method is preferable to the EWMAa method with respect to the average delay time. However, the predictive value of an early alarm is low for the EWMAe method. When considering also shifts at later time points, as is the natural choice in most applied situations, the picture is changed. In that case the EWMAa method will perform better, both with respect to the conditional expected delay, the expected delay and the probability of successful detection. This shows the importance to consider the performances of methods for shifts at different time points in every surveillance situation.

When comparing the detection ability of surveillance methods a common way is to fix the ARL^0 . Another approach is to compare methods for a fixed probability of a false alarm. Here, it is shown that when comparing the EWMAa and the EWMAe method with respect to the delay time of detecting a true change measured by the expected delay, this is a crucial choice. Choosing to fix the ARL^0 will favour the EWMAe method. If instead the probability of a false alarm is fixed, the EWMAa method performs relatively better.

There are several important differences between the one- and the two-sided versions of the methods. The most striking is the optimal value of λ for minimizing the ARL^1 for a fixed value of ARL^0 . For the one-sided case, λ should approach zero. This is true both for the EWMAa and the EWMAe method. However, this is not the case for the two-sided version of the EWMAa method where the optimal value of λ is larger, thus confirming results in Crowder (1989), Lucas and Saccucci (1990) and Srivastava and Wu (1997). This is somewhat surprising since the minimization of ARL^1 means minimizing the number of observations needed to distinguish between two possible distributions for all observations. To use different weights for different observations violates fundamental inference principles.

Another important difference concerns minimax-properties with respect to the delay time as a function of the time point of the shift. For the two-sided EWMAa method, the alarm limits will act in the same way as when using a one-sided version with barriers, and considerably alter the minimax properties of the method. This is especially apparent for small values of λ . The same pattern is not the case for the EWMAe method.

ACKNOWLEDGEMENT

The author wishes to thank Marianne Frisé for valuable support and comments during the work with this paper. Thanks also to Eva Andersson for valuable comments.

BIBLIOGRAPHY

- Borror, C.M., Montgomery, D.C., and Runger, G.C. (1999). "Robustness of the EWMA control chart to non-normality," *Journal of Quality Technology*, 31, 309-316.
- Box, G.E.P., Jenkins, G.M., and MacGregor, J.F. (1974). "Some recent Advances in Forecasting and Control," *Applied Statistics*, 23, 158-179.
- Champ, C.W. and Rigdon, S.E. (1991). "A Comparison of the Markov-Chain and the Integral-Equation Approaches For Evaluating the Run Length Distribution of Quality-Control Charts," *Communications in Statistics. Simulations and Computations*, 20, 191-204.
- Chan, L.K. and Zhang, J. (2000). "Some issues in the design of EWMA charts," *Communications in Statistics. Simulations and Computations*, 29, 207-217.
- Chandrasekaran, S., English, J.R., and Disney, R.L. (1995). "Modeling and Analysis of Ewma Control Schemes With Variance- Adjusted Control Limits," *IIE Transactions*, 27, 282-290.
- Chang, T.C. and Gan, F.F. (1994). "Optimal Designs of One-Sided Ewma Charts For Monitoring a Process Variance," *Journal of Statistical Computation and Simulation*, 49, 33-48.
- Crowder, S.V. (1987). "A simple method for studying run-length distributions of exponentially weighted moving average charts," *Technometrics*, 29, 401-407.
- Crowder, S.V. (1989). "Design of Exponentially Weighted Moving Average Schemes," *Journal of Quality Technology*, 21, 155-162.
- Crowder, S.V. and Hamilton, M.D. (1992). "An EWMA for monitoring a process standard-deviation," *Journal of Quality Technology*, 24, 12-21.
- Domangue, R. and Patch, S.C. (1991). "Some omnibus exponentially weighted moving average statistical process monitoring schemes.," *Technometrics*, 33, 299-313.
- Frisén, M. (1992). "Evaluations of Methods for Statistical Surveillance," *Statistics in Medicine*, 11, 1489-1502.
- Frisén, M. (1999). "Characterization of methods for surveillance by optimality," Research Report 1999:13 (revised 2000). Department of Statistics, Göteborg University
- Frisén, M. and de Maré, J. (1991). "Optimal Surveillance," *Biometrika*, 78, 271-80.
- Frisén, M. and Sonesson, C. (2001). "Optimal Surveillance with EWMA," *Manuscript*.
- Frisén, M. and Wessman, P. (1999). "Evaluations of likelihood ratio methods for surveillance. Differences and robustness.," *Communications in Statistics. Simulations and Computations*, 28, 597-622.
- Gan, F.F. (1993). "An optimal-design of EWMA control charts based on median run-length," *Journal of Statistical Computation and Simulation*, 45, 169-184.

- Gan, F.F. (1995). "Joint Monitoring of Process Mean and Variance Using Exponentially Weighted Moving Average Control Charts," *Technometrics*, 37, 446-453.
- Gan, F.F. (1998). "Designs of one- and two-sided exponential EWMA charts," *Journal of Quality Technology*, , 55-69.
- Järpe, E. (2000). On univariate and spatial surveillance. Unpublished PhD thesis. Department of Statistics, Göteborg university.
- Lorden, G. (1971). "Procedures for reacting to a change in distribution," *Annals of Mathematical Statistics*, 42, 1897-1908.
- Lu, C.W. and Reynolds, M.R. (1999). "EWMA control charts for monitoring the mean of autocorrelated processes," *Journal of Quality Technology*, 31, 166-188.
- Lucas, J.M. and Saccucci, M.S. (1990). "Exponentially weighted moving average control schemes: properties and enhancements," *Technometrics*, 32, 1-12.
- Morais, M.C. and Pacheco, A. (1998). "Two stochastic properties of one-sided exponentially weighted moving average control charts," *Communications in Statistics. Simulations and Computations*, 27, 937-952.
- Morais, M.C. and Pacheco, A. (2000). "On the performance of combined EWMA schemes for μ and σ : A Markovian approach," *Communications in Statistics. Simulations and Computations*, 29, 153-174.
- Moustakides, G.V. (1986). "Optimal stopping times for detecting changes in distributions," *Annals of Statistics*, , 1379-87.
- Pollak, M. (1985). "Optimal detection of a change in distribution," *Annals of Mathematical Statistics*, 13, 206-227.
- Roberts, S.W. (1959). "Control Chart Tests Based on Geometric Moving Averages," *Technometrics*, 1, 239-250.
- Robinson, P.B. and Ho, T.Y. (1978). "Average Run Lengths of Geometric Moving Average Charts by Numerical Methods," *Technometrics*, 20, 85-93.
- Runger, G.C. and Prabhu, S.S. (1996). "A Markov chain model for the multivariate exponentially weighted moving averages control chart," *Journal of the American Statistical Association*, 91, 1701-1706.
- Schöne, A., Schmid, W., and Knoth, S. (1999). "On the run length of the EWMA scheme: a monotonicity result for normal variables," *Journal of Statistical Planning and Inference*, 79, 289-297.
- Sonesson, C. and Bock, D. (2001). "Statistical issues in public health monitoring- A review and discussion," Research Report 2001:2. Department of Statistics, Göteborg University
- Srivastava, M.S. and Wu, Y. (1993). "Comparison of EWMA, CUSUM and Shiriyayev-Roberts Procedures for Detecting a Shift in the Mean.," *Annals of Statistics*, 21.

- Srivastava, M.S. and Wu, Y. (1997). "Evaluation of optimum weights and average run lengths in EWMA control schemes," *Communications in Statistics. Theory and Methods*, 26, 1253 - 1267.
- Steiner, S.H. (1999). "EWMA control charts with time-varying control limits and fast initial response," *Journal of Quality Technology*, 31, 75-86.
- Tsui, K.L. and Woodall, W.H. (1993). "Multivariate Control Charts Based on Loss Functions," *Sequential Analysis*, 12.

Research Report

- | | | |
|--------|---------------------------------------|--|
| 2000:1 | Hatemi-J, A. & Shukur, G.: | Multivariate based causality tests of twin deficits in the US. |
| 2000:2 | Petzold, M.: | Aspects on tolerance limit estimation - some common approaches and flexible modeling. |
| 2000:3 | Andersson, L.: | Statistical test of the existence of a turning point. |
| 2000:4 | Almasri, A.: | Some aspects of wavelet analysis in time series. |
| 2000:5 | Järpe, E.: | Surveillance of spatio-temporal patterns. Change of interaction in an Ising dynamic model. |
| 2000:6 | Järpe, E.: | Detection of environmental catastrophes. |
| 2001:1 | Holgersson, H.E.T.: | On assessing multivariate normality. |
| 2001:2 | Sonesson, C. & Bock, D.: | Statistical issues in public health monitoring – A review and discussion. |
| 2001:3 | Andersson, E.: | Turning point detection using non-parametric statistical surveillance. Evaluation of some influential factors. |
| 2001:4 | Andersson, E. & Bock, D.: | On seasonal filters and monotonicity. |
| 2001:5 | Andersson, E., Bock, D. & Frisén, M.: | Likelihood based methods for detection of turning points in business cycles. A comparative study. |