Control charts for monitoring the mean of AR(1) data

Wieringa, J.E.

Publication date:
1998

Control Charts for Monitoring the Mean of AR(1) Data

Jaap E. Wieringa*

March 12, 1998

Abstract

Control charts are designed to detect a special cause of variation in a sequence of observations. Traditionally, the assumption is made that successive observations are independently distributed. However, in practice, the observations are often serially correlated. This point has received considerable attention in the literature the last decade. Roughly speaking, two kinds of control charts were developed in order to deal with serial correlation in the observations: the modified Shewhart chart and the residuals chart. In this paper it is investigated how well these control charts are able to detect a shift in the mean of AR(1) data. The performance of these charts is compared to the well studied case of independent observations. It turns out that for negative autocorrelation, the residuals chart is the best of these two, for positive autocorrelation it is best to choose the modified Shewhart chart. A modification of the residuals chart is developed that outperforms both charts in case of positive autocorrelation, which is most commonly encountered in practice. In case of negative autocorrelation, the residuals chart remains the best choice. The procedure is illustrated by two examples. The first is based on a data set from Shewhart (1931), the second is a simulated example.

Keywords: Residual Control Charts, SPC, Time Series models.

1 Introduction

Control charts were originally developed by Walter A. Shewhart in the twenties, and are widely used in practice as a tool for reducing variability in process outcomes. The success of the control chart is based on Shewhart's classification of sources of variation into two groups: common causes of variation,

*Department of Econometrics, Faculty of Economic Sciences, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands. E-mail j.e.wieringa@eco.rug.nl

and special causes of variation. The variation due to common causes of variation is the joint effect of numerous small causes that act upon the outcomes of the process, and that cannot be removed without a profound revision of the process. This variation is inherently part of the process, and is always present. Usually, it is not within the power of an operator to influence the effect of common causes of variation on product outcomes. The second group of sources of variation consists of *special causes of variation*. They are *not* part of the process, and occur only accidently. However, when a special cause of variation is present, it will have a large effect on process outcomes. If removal is possible, a special cause can usually be eliminated without revising the process. In many cases, an operator can be instructed to recognize and remove special causes of variation, thereby improving the quality of the outcomes of the process. It is the task of the control chart to discriminate between situations where only common causes are affecting the outcomes of a process and situations where there are also special causes of variation present. If only common causes are present, the process is said to be *in control*.

In most literature on SPC it is assumed that X_t, an observation of a quality characteristic at time t, can be modelled as $X_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu_t, \sigma^2_{X_t})$, for $t \in \mathbb{Z}$. The index t in μ_t and $\sigma^2_{X_t}$ indicates that the mean and the variance of the observations may change in time due to the presence of special causes of variation. With these assumptions, the process is in control if and only if $\mu_t = \mu$ and $\sigma^2_{X_t} = \sigma^2_X$ for all t for certain μ and σ^2_X. For this reason, a production process is usually monitored using two control charts, one for monitoring the variance, and one for monitoring the mean of the process. Control charts that can be used include the Shewhart, the CUSUM, and the EWMA control chart (see for example textbooks like Ryan (1989) or Montgomery (1996)).

In this paper, we will limit ourselves to Shewhart type control charts for the mean. Therefore, we make the assumption that a special cause can only affect the mean of the process, and that the variance is constant. The statistical properties of a Shewhart type control chart for detecting a shift in the mean of independent observations are summarized in Section 2.

In the sections following Section 2, the independence assumption is loosened to include observations from an AR(1) process. For detecting a shift in the mean of such a process two approaches are suggested in the literature. The first approach is a modification of the classical Shewhart control chart. The control limits are modified to allow for serial correlation in the data. The modified Shewhart chart is discussed in Section 3. The second approach is a so-called residuals control chart. It will be discussed in Section 4. On the basis of Average Run Length (ARL) curves, the performance of these charts is compared to the performance of the Shewhart chart for independent observations. From this comparison, it is possible to evaluate the loss of efficiency due to serial correlation.

In Section 5, a modification of the residuals control chart is presented that has a better ARL performance than the modified Shewhart and the residuals control chart.

A simulated example is discussed in Section 6, and the conclusions are summarized in Section 7.

2 The i.i.d. case

In the classical situation, it is assumed that subsequent observations of a quality characteristic are independently distributed. In this paper, we will denote a sequence of independent observations by $\{X_t\}$. An observation may be the mean of a sample, or an individual observation. We will assume that the variance of the observations remains constant over time, and that a special cause of variation may cause a shift in the mean of the process. Assuming normality, we have the following model for X_t, an observation of the quality characteristic at time t

$$X_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu_t, \sigma_X^2) \qquad \text{for } \ t \in \mathbb{Z}, \tag{2.1}$$

where the expectation of X_t is indexed by time, to indicate that the mean of the process may shift due to special causes of variation.

A special cause occurring at an unknown time point T is modelled as

$$\mu_t = \begin{cases} \mu & \text{for } \ t < T \\ \mu + \delta\sigma_X & \text{for } \ t \geq T. \end{cases} \tag{2.2}$$

The size of the shift is expressed in units of the standard deviation of the observations. This facilitates comparison of control charts for processes with different variances.

The classical Shewhart control chart has control limits $\mu \pm 3\sigma_X$. After a shift of size $\delta\sigma_X$ has occurred, we can write for $\mathrm{P}(\delta)$, the probability that an observation falls within the control limits

$$\begin{aligned} \mathrm{P}(\delta) &= \mathrm{P}(\mu - 3\sigma_X \leq X_t \leq \mu + 3\sigma_X) \\ &= \Phi(\delta + 3) - \Phi(\delta - 3), \end{aligned}$$

where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution. It can be computed that $\mathrm{P}(0) = 0.9973$, so that it is very unlikely to observe an out of control signal if $\delta = 0$. Therefore, if an observation outside the control limits is encountered, the presence of a special cause of variation is suspected. This interpretation of an out of control signal relies on the independence and normality assumptions made for model (2.1). In the following sections, the independence assumption is loosened to include observations from an AR(1) model. The performance of control charts for AR(1) models is compared to the performance of the Shewhart chart for independent observations on the basis of ARL curves.

The ARL is defined as the average number of observations upto and including the first out of control observation. The ARL is a function of δ. For the case of independent observations, $\mathrm{ARL}(\delta)$ can be computed as

$$\mathrm{ARL}(\delta) \quad = \quad \sum_{i=1}^{\infty} i\mathrm{P}(\delta)^{i-1}[1-\mathrm{P}(\delta)] \quad = \quad \frac{1}{1-\mathrm{P}(\delta)}.$$

In Figure 2.1, the ARL curve of a Shewhart chart with three sigma limits is depicted.

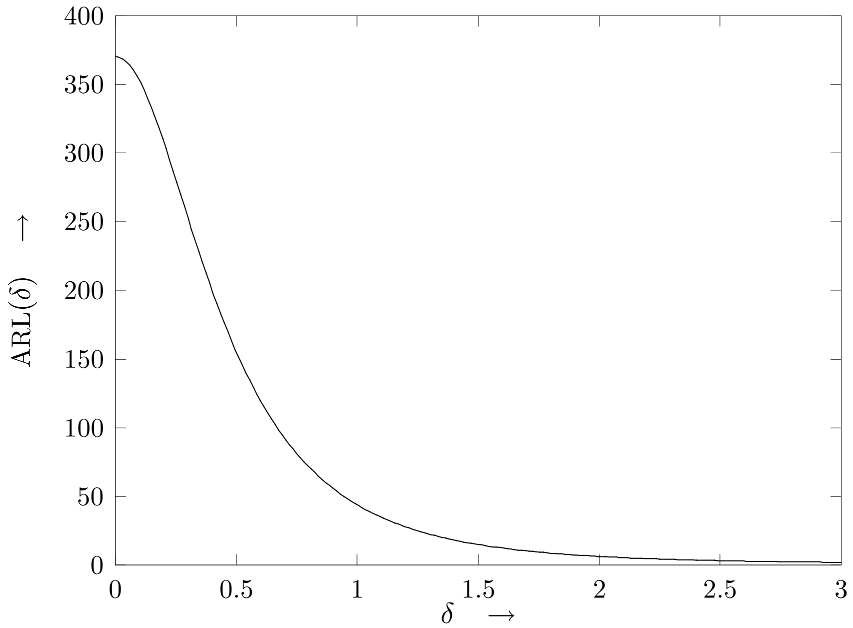

Figure 2.1: ARL curve of a Shewhart chart.

Figure 2.1 shows that the ARL is high if $\delta = 0$, and that the ARL is low if δ is large. This is desirable behavior. The ARL curves of control charts discussed in the remainder of this paper will be compared to the curve in

Figure 2.1. A control chart having the same ARL(0), and lower ARL(δ) for $\delta > 0$ is more efficient, since on average less observations are needed to detect a change in $\mathrm{E}(X_t)$. Analogously, a control chart having the same ARL(0) and higher ARL(δ) for $\delta > 0$ is less efficient.

From a theoretical point of view it is interesting to investigate the effect of serial correlation on the performance of control charts. However, the main motivation for this extension stems from practical situations, where the assumption of independence of the observations is often not fulfilled. Serial correlation in observations may for example arise from mixing of tanks with raw material, so that it is not always possible to remove serial correlation without affecting the proper functioning of the process. It has been the experience of many SPC practitioners that simply ignoring serial correlation by monitoring the process with a control chart designed for independent observations produces misleading results. In the next section, it is shown in figure 3.1 that first order autocorrelation makes a control chart less sensitive for changes in the mean. When designing a control chart for serially correlated data, the correlation should be accounted for.

3 The modified Shewhart control chart

3.1 The AR(1) process

In the following sections, we will evaluate the ability of control charts to detect a shift in the mean of an AR(1) process. In this subsection, we will briefly describe the AR(1) process. To distinguish from a sequence of independent observations, a sequence of AR(1) observations will be denoted by $\{Y_t\}$. Again, an observation may be the mean of a sample or an individual observation. We assume that Y_t is generated by a stationary AR(1) model if the process is in-control:

$$Y_t - \mu = \phi(Y_{t-1} - \mu) + \varepsilon_t \qquad \text{for } t \in \mathbb{Z}, \tag{3.1}$$

where $\{\varepsilon_t\}$ is a sequence of i.i.d. disturbances, $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$ for $t \in \mathbb{Z}$. Subsequent observations of model (3.1) are serially correlated, since

$$\mathrm{Cov}(Y_t, Y_{t-k}) = \phi^k \sigma_Y^2,$$

where

$$\sigma_Y^2 = \mathrm{Var}(Y_t) = \frac{\sigma^2}{1 - \phi^2}.$$

Obviously, $\mathrm{E}(Y_t) = \mu$ for all t in model (3.1). A model that includes the possibility of a shift in $\mathrm{E}(Y_t) \equiv \mu_t$ is

$$Y_t - \mu_t = \phi(Y_{t-1} - \mu_{t-1}) + \varepsilon_t \qquad \text{for } t \in \mathbb{Z}. \tag{3.2}$$

The process is in-control if $\mu_t = \mu$ for all t in $\mathbb{Z}$. Analogous to the i.i.d. case, we want to study the ARL behavior of a control chart that is designed to detect a shift of size $\delta\sigma_Y$ in $\mathrm{E}(Y_t)$. Note that the i.i.d. case is a special case of the process discussed in this section, since if $\phi = 0$, model (3.2) reduces to model (2.1).

Several authors have considered detecting a shift in model (3.2). Their work can roughly be classified into two categories. The first group of authors recommends plotting the original observations, and adjust the control charts for the effect of serial correlation that is present in the data. This approach will be discussed in this section. The second group suggests fitting a time series model, and monitor the residuals for a departure from the in-control model. This will be discussed in the next section.

3.2 Modified Shewhart limits

The first group of authors recommends plotting the original observations in control charts with modified limits. These authors include Vasilopoulos and Stamboulis (1978) who discuss modifications of Shewhart control charts to account for serial correlation, Johnson and Bagshaw (1974) and Bagshaw and Johnson (1975), who investigated the effect of serial correlation on CUSUM control charts. Schmid (1997) proposes a modification of the EWMA control chart to account for serial correlation.

The word 'modified' refers to two adaptations compared to the control chart for i.i.d. observations. The first adaptation is very natural: for determining the width of the control limits the variance of the observations is used instead of the variance of the disturbances. The variance of Y_t is σ^2, multiplied with a factor of $1/(1-\phi^2) > 1$. If ϕ is in- or decreased from zero, the variance of the observations increases. A shift in the mean that is hidden in a sequence of observations with larger variance is harder to detect. Therefore, it is to be expected that monitoring AR(1) observations driven by white noise with variance σ^2 for a shift in the mean of size $\delta\sigma$ is not as efficient as monitoring an independent series with variance σ^2 for a shift of the same size. The first adaptation of the control limits compensates for the increase in the variance of the observations: the observations from model (3.2) are compared to control limits $\mu \pm 3\sigma_Y$ instead of $\mu \pm 3\sigma$.

The second adaptation follows from the ARL properties of the first adaptation. Hence, it will be discussed after we have evaluated the ARL.

For evaluating the ARL of the modified Shewhart chart, the Markov chain approach can be used (see Brook and Evans (1972)). However, we will use the integral equation approach presented by Crowder (1987). Suppose that the sequence $\{Y_t\}$ is generated by (3.2), and that on time T a shift of size $\delta\sigma_Y$ has occurred in $\mathrm{E}(Y_t)$. For $t \geq T+1$, the observations of Y_t are then generated by

$$(Y_t - \mu) - \delta\sigma_Y = \phi\,(Y_{t-1} - \mu) - \phi\delta\sigma_Y + \varepsilon_t$$

If the value of Y_{t-1} is s, we can write for v, the realization of Y_t

$$\begin{aligned} v &= \mu + \delta\sigma_Y + \phi\,(s - (\mu + \delta\sigma_Y)) + \varepsilon_t \\ &= \phi s + (1-\phi)\mu + (1-\phi)\delta\sigma_Y + \varepsilon_t, \end{aligned}$$

provided $t \geq T+1$. The observations are compared to control limits $\mu \pm 3\sigma_Y$. The run length is one if the next observation v falls outside the control limits. If v is within the control limits, the run length is one *plus* some additional observations, which can be regarded as a run length of the AR(1) process, starting in v. Let $L_\phi(\delta, s)$ denote the ARL of the modified Shewhart control chart as a function of δ and s, the value of the AR(1) process when we started to take observations. The latter is of importance since the observations are serially correlated. The function also depends on the AR parameter ϕ. We can write the following for $L_\phi(\delta, s)$:

$$\begin{aligned} L_\phi(\delta, s) &= 1 + \int_{\{\varepsilon \mid \mu - 3\sigma_Y \leq v \leq \mu + 3\sigma_Y\}} L_\phi\,(\delta, v)\, f(\varepsilon)\, \mathrm{d}\varepsilon \\ &= 1 + \int_{\mu - 3\sigma_Y}^{\mu + 3\sigma_Y} L_\phi\,(\delta, v)\, f\,[v - \phi s - (1-\phi)(\mu + \delta\sigma_Y)]\, \mathrm{d}v, \end{aligned}$$

where $f(\varepsilon)$ is the density function of ε. Note that for $\phi = 0$, $L_\phi(\delta, s)$ does not depend on s, and hence is a constant relative to s. This makes sense because the observations are independent if $\phi = 0$. From the integral equation above it follows that $L_0(\delta, s) = \mathrm{ARL}(\delta)$ for all $s \in (\mu - 3\sigma_Y, \mu + 3\sigma_Y)$, as expected. The unknown function $L_\phi(\delta, s)$ can be numerically evaluated using Gaussian Quadrature.

The function $L_\phi(\delta, s)$, is the ARL function of the modified Shewhart chart if we start to take observations *after* the shift has occurred. In practice, we are often interested in how quickly on average a shift is detected starting from the moment the shift occurs. The computation of this ARL function is slightly different. If it is assumed that the value of Y_{T-1} is s, we can write for v^*, the realization of Y_T

$$
\begin{aligned}
v^* &= \mu + \delta\sigma_Y + \phi\,(s - \mu) + \varepsilon_T \\
&= \phi s + (1-\phi)\mu + \delta\sigma_Y + \varepsilon_T.
\end{aligned}
$$

We will denote the ARL of a modified Shewhart chart when the first observation is taken at the time of the shift by $L^*_\phi(\delta, s)$. This function is related to $L_\phi(\delta, s)$ as follows

$$
L^*_\phi(\delta, s) \;=\; 1 \;+\; \int_{\mu-3\sigma_Y}^{\mu+3\sigma_Y} L_\phi\,(\delta, v^*)\, f[v^* - \phi s - (1-\phi)\mu + \delta\sigma_Y]\; \mathrm{d}v^*.
$$

The difference between $L^*_\phi(\delta, s)$ and $L_\phi(\delta, s)$ is small. For negative values of ϕ, $L^*_\phi(\delta, s) \geq L_\phi(\delta, s)$, while for positive ϕ, $L^*_\phi(\delta, s) \leq L_\phi(\delta, s)$.

In Figure 3.1 $L^*_\phi(\delta, 0)$ curves are drawn for $\phi = -0.6, 0, 0.6$ as functions of δ. The curve for $\phi = 0$ is identical to the curve depicted in Figure 2.1.

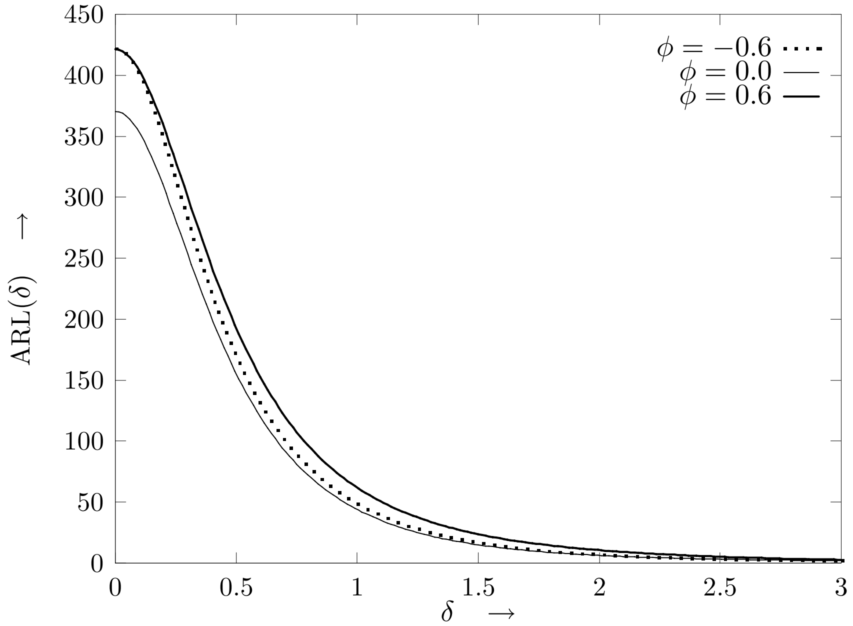

Figure 3.1: ARL curves of the modified Shewhart chart for various AR(1) processes.

The ARL curves in figure 3.1 show the result of a simple adaptation of the Shewhart chart. The correlation is basically ignored. The width of the control limits is based on σ_Y, which is the standard deviation of the observations, and not on σ, the standard deviation of the disturbances. Figure 3.1 shows that, compared to the Shewhart chart for i.i.d. observations, the adapted control chart has a higher in-control ARL. This agrees with a result in Schmid (1995), where it was proven that in-control ARL values for arbitrary Gaussian processes are always larger than the in-control ARL for i.i.d. observations. This effect is advantageous. For $\delta > 0$, the ARL curves are adversely affected by first order autocorrelation. Compared to the i.i.d. case, the adapted control chart is less sensitive for detecting shifts in the mean.

The curves drawn in Figure 3.1 show the typical behavior of $L^*_\phi(\delta, 0)$ for nonzero ϕ in the sense that a nonzero ϕ results in a larger $L^*_\phi(0, 0)$, but also in larger $L^*_\phi(\delta, 0)$ for $\delta > 0$.

Whether the net effect of first order autocorrelation is beneficial or not is not clear from Figure 3.1. To make a proper evaluation, a second adaptation of the control chart is needed. The adaptation consists of tightening the limits of the adapted Shewhart chart in such a way that $L^*_\phi(0, 0) \approx \text{ARL}(0)$.

The resulting ARL curves for $\phi = -0.9$ and $\phi = 0$ are depicted in Figure 3.2. ARL curves of intermediate values of ϕ were not drawn, since there is no visible difference with the curve of $\phi = 0$ (this can be checked in Figures 4.2 and 4.3). The curves for various $\phi \geq 0$ are drawn in Figure 3.3.

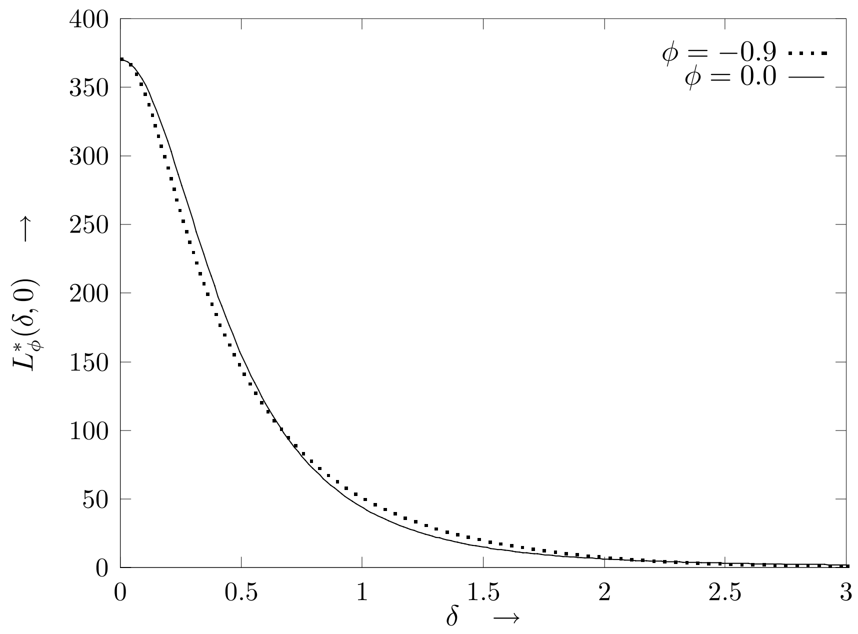

Figure 3.2: $L^*_\phi(\delta, 0)$ as a function of δ for $\phi = -0.9$ and $\phi = 0$.

Figures 3.2 and 3.3 were drawn with the objective to show that all but one ARL curves are close together. The ARL curves corresponding to negative

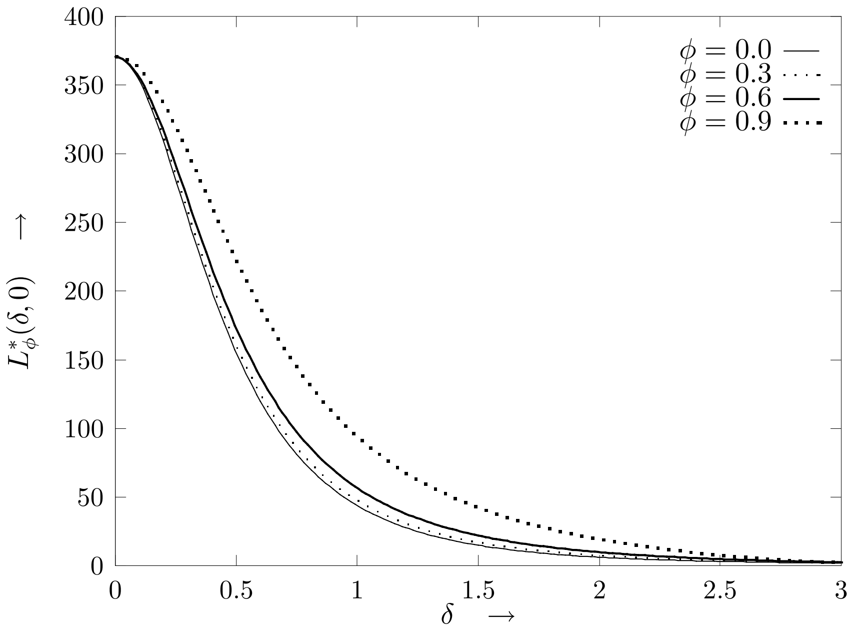

Figure 3.3: $L^*_\phi(\delta, 0)$ as a function of δ for various $\phi \geq 0$.

values of ϕ practically coincide with the curve corresponding to the i.i.d. case. For positive values of ϕ, the modified Shewhart chart is performing worse. However, for $\phi = 0.3$ and $\phi = 0.6$ the differences with the i.i.d. case are small. For the values of ϕ considered, the modified Shewhart chart behaves only considerably worse for $\phi = 0.9$.

In conclusion, if one monitors AR(1) data for a shift in the mean with the aid of a modified Shewhart control chart, then the performance is comparable to a Shewhart chart for independent observations, provided ϕ is not too large.

3.3 Discussion

In this subsection the modified Shewhart chart for AR(1) data is compared to the EWMA control chart for independent observations. It turns out that they are very similar. The EWMA control chart was firstly presented by Roberts (1959). More recent references include Hunter (1986), Crowder (1987), and Lucas and Saccucci (1990). In these references, it is shown that the ARL behavior of the EWMA control chart is better than the ARL performance of the Shewhart chart for independent observations. Therefore, one might expect that the ARL performance of the modified Shewhart chart for AR(1) data is also better than that of the Shewhart chart for independent observations. However, in the previous subsection it was argued that the ARL performance of the modified Shewhart chart for AR(1) data is comparable to that of the Shewhart chart for independent observations. In this subsection, it is explained why the ARL performance of the modified Shewhart chart for AR(1) data is not as good as the ARL performance of the EWMA chart for independent observations.

Suppose that we have a sequence of i.i.d. observations $\{X_t\}$, which satisfy (2.1) and (2.2). The EWMA statistic at time t will be denoted by W_t and is constructed as follows

$$W_t = (1-\lambda)W_{t-1} + \lambda X_t \qquad \text{for } t = 1, 2, \cdots, \tag{3.3}$$

where $\lambda \in (0,1)$. The EWMA chart may be started by setting W_0 equal to a target value or (an estimator of) μ. If $W_0 = \mu$, it is easy to verify that $\mathrm{E}(W_t) = \mu$ for $t = 0, 1, 2, \cdots$,T. For $t \geq T$ we have that

$$\mathrm{E}(W_t) = \mu + \left[1 - (1-\lambda)^{t-T+1}\right] \delta\sigma_X,$$

which approximately equals $\mu + \delta\sigma_X$ for $t \gg T$. Hence, $\{\mathrm{E}(W_t)\}$, the sequence of expected values of the EWMA statistic, approximately mimics $\{\mathrm{E}(X_t)\}$.

The relation between (3.3) and the AR(1) model (3.2) becomes prevalent if we subtract μ_t on both sides of equation (3.3):

$$W_t - \mu_t = (1-\lambda)(W_{t-1} - \mu_t) + \lambda(X_t - \mu_t). \tag{3.4}$$

Note that $\{\lambda(X_t - \mu_t)\}$ may be considered a white noise process with variance $\lambda^2\sigma_X^2$. As long as $\mu_t = \mu$ for $t = 0, 1, 2, \cdots$, we conclude that computing the EWMA statistic is equivalent to converting a sequence of i.i.d. observations into an AR(1) sequence with AR parameter $\phi = 1 - \lambda$. Moreover, using an EWMA control chart is in some sense equivalent to monitoring AR(1) observations in a Shewhart type control chart. In Lucas and Saccucci (1990) it was shown that the properties of the EWMA are very close to those of CUSUM schemes. Small shifts in the mean are on average more quickly detected on an EWMA control chart than on a standard Shewhart control scheme.

The argument above seems somewhat in contradiction with the conclusions of the previous subsection. There we observed that for $\phi = 0.9$ the modified Shewhart control chart for AR(1) data is not very sensitive in detecting small changes in the mean. A value $\phi = 0.9$ corresponds to a value of $\lambda = 0.1$. For this value of λ, an EWMA chart is much more sensitive than a Shewhart control chart for detecting small shifts in the mean of a sequence of independent observations. The discrepancy between these results can be explained by studying a 'signal to noise' ratio: a number that relates the size of the shift to the standard deviation of the process. This ratio allows us to compare shifts in means in processes with different variances. We define the 'signal to noise' ratio as the size of the shift divided by the standard deviation of the process.

For the modified Shewhart chart, the size of the shift is $\delta\sigma_Y$ after time T. The standard deviation of the AR(1) process is σ_Y. Hence, the 'signal to noise' ratio is δ.

In the case of the EWMA control chart, the size of the shift in $\mathrm{E}(W_t)$ approximately equals $\delta\sigma_X$ for $t \gg T$. The variance of the EWMA statistic is

$$\mathrm{Var}(W_t) = \sigma_X^2 \left(\frac{\lambda}{2-\lambda}\right)\left[1-(1-\lambda)^{2t}\right].$$

Hence, for large $t \gg T$, the 'signal to noise' ratio equals approximately $\delta\sqrt{(2-\lambda)/\lambda}$, which is larger than δ for $0 < \lambda < 1$. A popular choice of λ for the EWMA chart is $\lambda = 0.1$. In this case, the 'signal to noise' ratio is approximately $\delta\sqrt{19}$ for large $t \gg T$.

Hence, computing the EWMA of a sequence of i.i.d. observations leaves the pattern of expectations approximately unaltered, but improves the 'signal to noise' ratio. This is combined with the introduction of first order positive autocorrelation. From the ARL curves in the previous subsection, we learned that first order positive autocorrelation has a negative effect on the performance of the ARL. Apparently, this effect is offset by the positive effect of the improved 'signal to noise' ratio. We conclude that the efficiency of the EWMA control chart is not the result of the autocorrelation that is introduced. It is the improvement of the 'signal to noise' ratio that makes the EWMA control chart efficient.

4 The residuals control chart

A second approach for monitoring AR(1) observations is amongst others discussed by Montgomery and Mastrangelo (1991), Alwan and Roberts (1988), Berthouex, Hunter and Pallesen (1978), Harris and Roberts (1991), and in Kramer and Schmid (1996). The idea is to monitor the process using residuals of fitting a time series model to the data. If a shift in the mean of the process occurs, this will cause a shift in the mean of the residuals. Furthermore, if the time series model fits the data well, the residuals will be approximately uncorrelated. This provides a theoretically elegant way to monitor a serially correlated process using control charts that were designed for independent observations.

4.1 Residuals of an AR(1) process

The residual control chart is based on charting residuals

$$e_t \equiv Y_t - \hat{Y}_{t|t-1,t-2,\cdots}, \tag{4.1}$$

where $\hat{Y}_{t|t-1,t-2,\cdots}$ is a forecast of Y_t, based on observations upto and including time $t-1$. The linear forecast that minimizes the mean square error is $\mathrm{E}(Y_t|Y_{t-1}, Y_{t-2}, \cdots)$, see for example Harvey (1993). In the case of AR(1) data generated by the in-control model (3.1), we have

$$\hat{Y}_{t|t-1,t-2,\cdots} = \mathrm{E}\left(Y_t|Y_{t-1}, Y_{t-2}, \cdots\right) = \mu + \phi(Y_{t-1} - \mu). \tag{4.2}$$

In practice, μ and ϕ will have to be estimated from a data set that was obtained in a period where only common causes of variation were affecting the process. Throughout this paper we will assume that enough in-control observations are available so that μ and ϕ can be estimated accurately.

As long as the process is in control, observations are generated by model (3.1), and e_t, the quantities that will be plotted in the residuals control chart satisfy

$$e_t = Y_t - \hat{Y}_{t|t-1,t-2,\cdots} \doteq \varepsilon_t \qquad \text{for all } t \tag{4.3}$$

where the last relation is exact if μ and ϕ are known. Suppose that a special cause shifts $\mathrm{E}(Y_t)$ at time T by an amount of $\delta\sigma_Y$. Since we are not aware of this shift, we compute $e_T, e_{T+1}, \cdots$ as if the process were in control. Hence, also for $t > T$, computation of e_t is given by (4.1), with $\hat{Y}_{t|t-1,t-2,\cdots}$ computed as in (4.2).

The elements of the sequence of residuals $\{e_t\}$ satisfy

$$e_t \doteq \begin{cases} \varepsilon_t & \text{for } t < T \\ \varepsilon_t + \delta\sigma_Y & \text{for } t = T \\ \varepsilon_t + (1-\phi)\delta\sigma_Y & \text{for } t = T+1, T+2, \cdots. \end{cases} \tag{4.4}$$

Note that the sequence $\{e_t\}$ is approximately independently distributed since we assumed that $\varepsilon_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma^2)$. Hence, we are back at the i.i.d. case. For $\phi > 0$ only a fraction of the shift in $\mathrm{E}(Y_t)$ is transferred to the residuals for $t > T$. For $\phi < 0$, the shift is blown up. Therefore, we expect the residuals chart to perform better for AR(1) data with negative ϕ relative to AR(1) data with positive ϕ, since the 'signal to noise' ratio is higher for negative ϕ.

4.2 The ARL of the residuals chart

In the case where the first observation is taken after the shift in $\mathrm{E}(Y_t)$ has occurred, the computation of the ARL of the residuals chart is analogous to the computation of the ARL curve of a Shewhart chart for independent observations, as described in Section 2.

However, if the first observation is taken at the time of the shift, or if we are interested in the expected number of observations that a shift goes undetected, the computation is slightly different. The reason for this is that the probability of observing a residual between the the control limits at the time of the shift differs from the probability that a residual falls between the control limits after the shift. Let us denote the probability that a residual falls between the control limits for $t > T$ by $\mathrm{P}(\delta)$, and let $\mathrm{P}_1(\delta)$ denote the probability that a residual falls between the limits at time T. It can be shown that $\mathrm{ARL}_{\mathrm{rc}}(\delta)$, the ARL of the residuals chart, satisfies

$$\mathrm{ARL}_{\mathrm{rc}}(\delta) = 1 + \frac{1}{1 - \mathrm{P}(\delta)} \mathrm{P}_1(\delta) \tag{4.5}$$

if the first observation is taken at the time of the shift. Note that, if $\mathrm{P}_1(\delta) = \mathrm{P}(\delta)$, the right hand side reduces to $1/(1 - \mathrm{P}(\delta))$, which is the ARL of the residuals chart if the first observation is taken after the shift in $\mathrm{E}(Y_t)$ has occurred.

The difference between $\mathrm{ARL}_{\mathrm{rc}}(\delta)$ and $1/(1-\mathrm{P}(\delta))$ is negligible for negative ϕ. However for large positive ϕ the difference is quite large. This can be explained by looking at 'signal to noise' ratios. At time T, the size of the shift that is transferred to the residuals is approximately equal to $\delta\sigma_Y$. Dividing this quantity by σ, we have that

$$\frac{\delta\sigma_Y}{\sigma} = \frac{\delta}{\sqrt{1 - \phi^2}}.$$

Hence, the 'signal to noise' ratio converges to infinity as $\phi \to 1$. As a result, the probability that the shift will be detected at the first observation converges to one as $\phi \to 1$. Consequently, $\mathrm{P}_1(\delta) \to 0$ for $\delta \neq 0$. This affects $\mathrm{ARL}_{\mathrm{rc}}(\delta)$ positively, since it follows from equation (4.5) that $\mathrm{ARL}_{\mathrm{rc}}(\delta)$ converges to 1 as $\mathrm{P}_1(\delta) \to 0$.

For $t > T$, the 'signal to noise' ratio is approximately equal to

$$\frac{(1 - \phi)\delta\sigma_Y}{\sigma} = \delta \, \frac{\sqrt{1 - \phi}}{\sqrt{1 + \phi}}, \tag{4.6}$$

which converges to 0 as $\phi \to 1$. Hence, for values of ϕ very close to one it is very hard to detect a shift if it is not detected at the first observation.

In Figures 4.1 through 4.6 ARL curves of the residuals chart are compared to ARL curves of the modified Shewhart chart for various values of ϕ. For these curves it is assumed that the first observation is taken at the time of the shift. As a reference, also the ARL curve for the i.i.d. case is depicted.

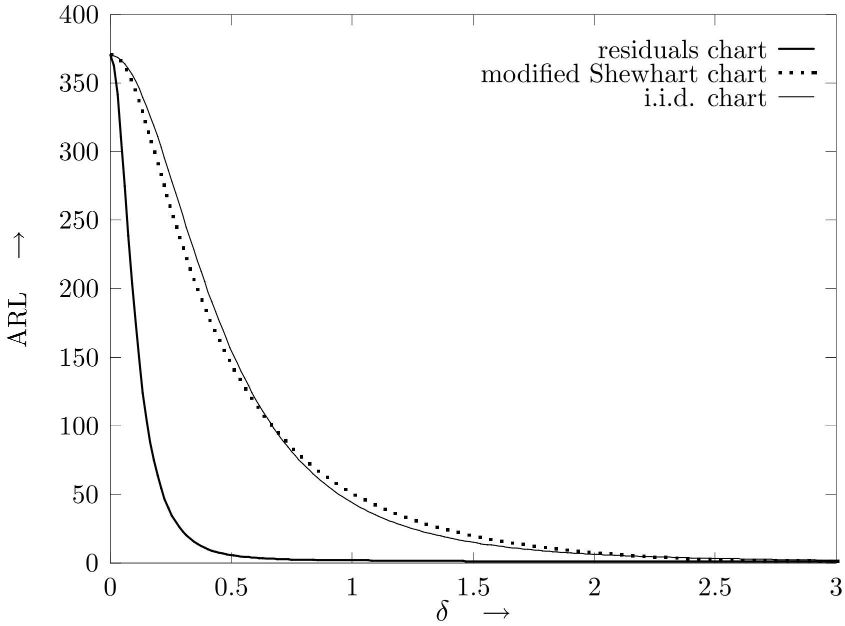

Figure 4.1: Various ARL curves for AR(1) process with $\phi = -0.9$.

The ARL curves for the modified Shewhart chart were depicted earlier in Figures 3.2 and 3.3. In these figures, also the ARL curve for $\phi = 0$ was depicted. Since all three curves coincide for $\phi = 0$ (the i.i.d. situation), we did not include a graph for this case here.

4.3 Discussion

From Figures 4.1 through 4.6, we conclude that for negative first order autocorrelation, the residuals chart is performing better than the Shewhart chart for independent observations. This was to be expected, since in Subsection 4.1 it was shown that for negative autocorrelation, a shift in AR(1) observations is blown up in the residuals.

Compared to the Shewhart chart for independent observations, the performance of the residuals chart is worse for positive autocorrelation. This

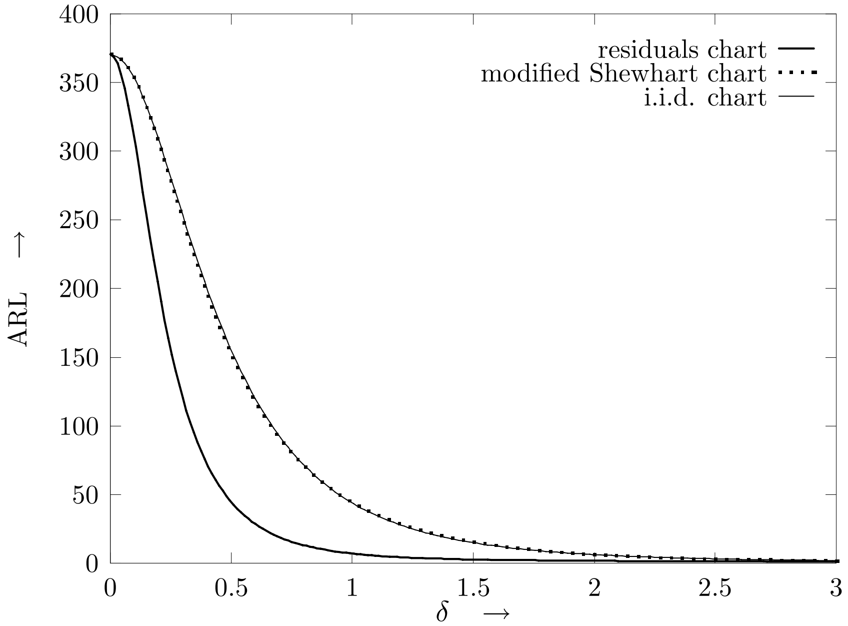

Figure 4.2: Various ARL curves for AR(1) process with $\phi = -0.6$.

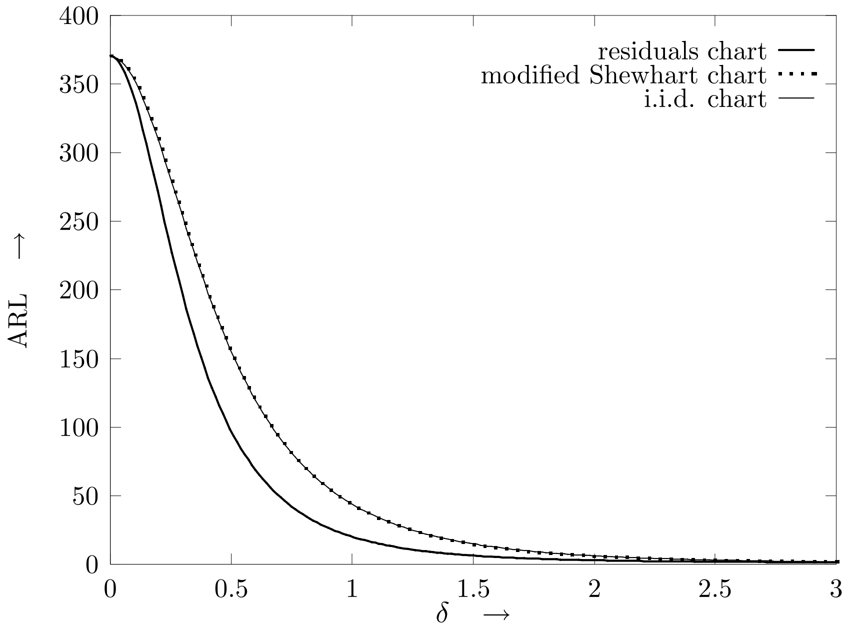

Figure 4.3: Various ARL curves for AR(1) process with $\phi = -0.3$.

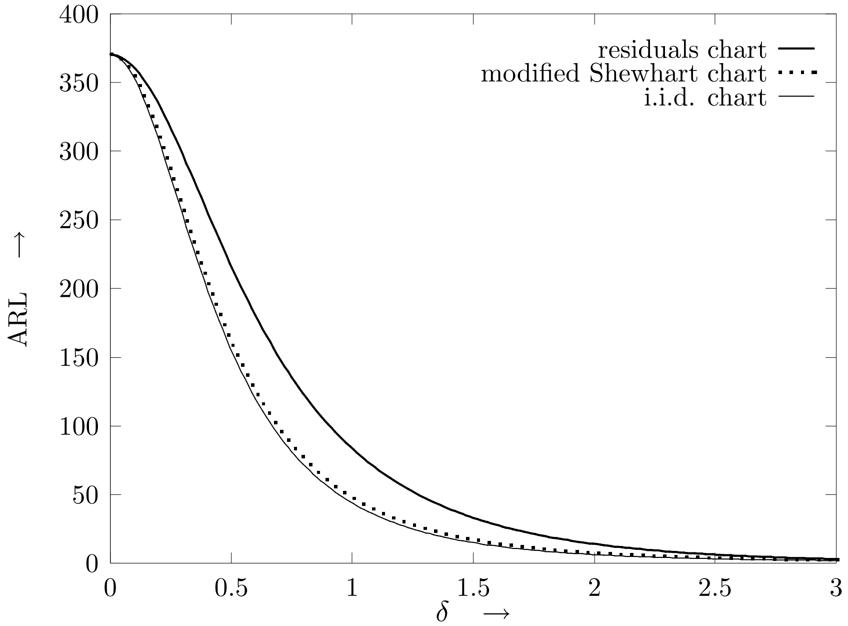

Figure 4.4: Various ARL curves for AR(1) process with $\phi = 0.3$.

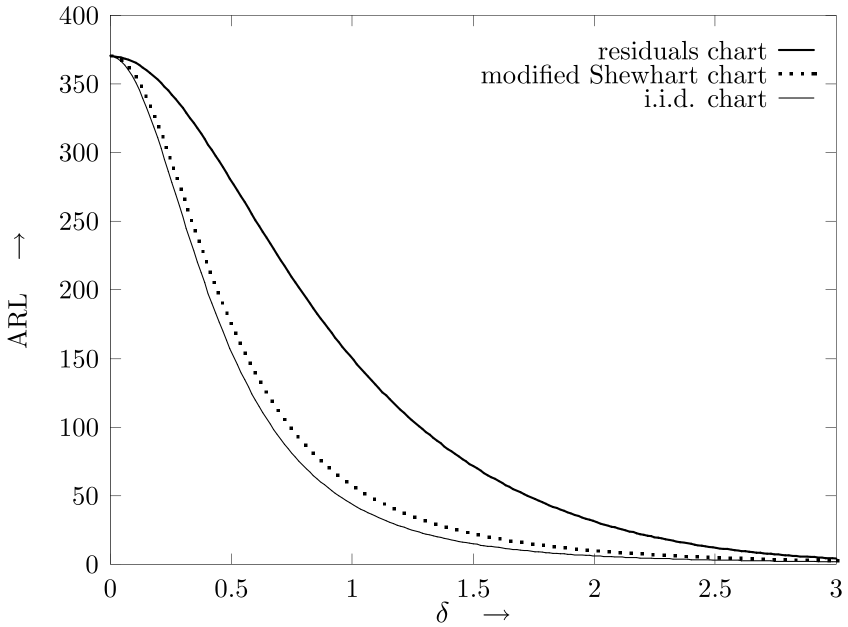

Figure 4.5: Various ARL curves for AR(1) process with $\phi = 0.6$.

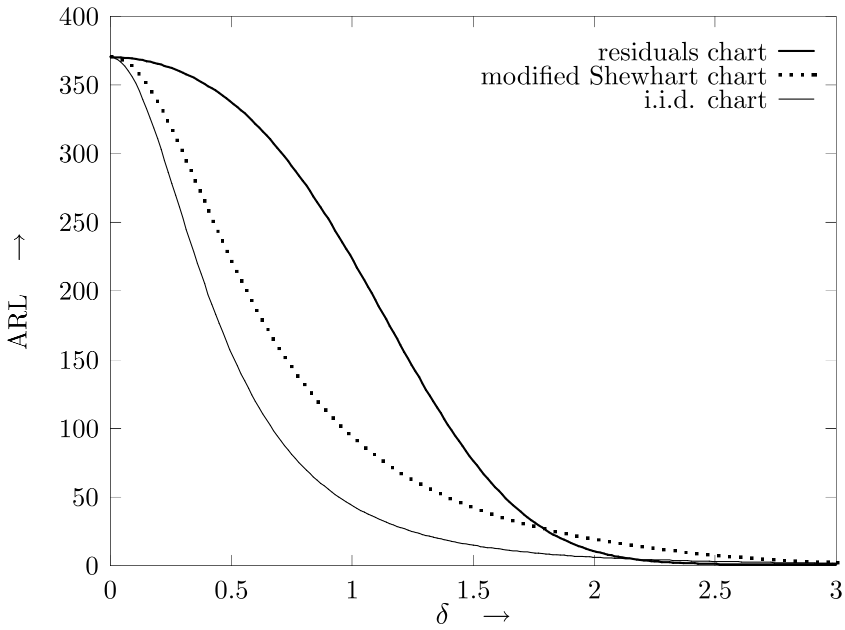

Figure 4.6: Various ARL curves of AR(1) process with $\phi = 0.9$.

is caused by the fact that only a fraction of the shift is transferred to the residuals for positive ϕ.

For values of ϕ very close to one, we argued in the previous section that the corresponding ARL curve will converge to 1 for all $\delta > 0$. A value of $\phi = 0.9$ is not close enough to 1 to show this convincingly in Figure 4.6. Therefore, in Figure 4.7, the three ARL curves are drawn for $\phi = 0.99$. This graph shows that the ARL performance of the residuals chart is improved for values of ϕ very close to 1.

The conclusions drawn from Figures 4.1 through 4.6 and Figure 4.7 agree with a comment of Ryan (1991):

> "*A residuals chart for* AR(1) *data will perform poorly unless* ϕ *is negative or extremely close to* 1. *In most applications we would expect to have* $\hat{\phi} > 0$ *and not particularly close to one*".

This is an important disadvantage of the residuals chart.

At the start, the residuals chart seemed to be attractive. By removing the serial correlation, the problem is reduced to the well known case of detecting a shift in the mean of independent observations. In contrast, the modified Shewhart chart basically ignores the serial correlation in the data. The control limits are adjusted in a rather ad-hoc manner to ensure a certain in-control

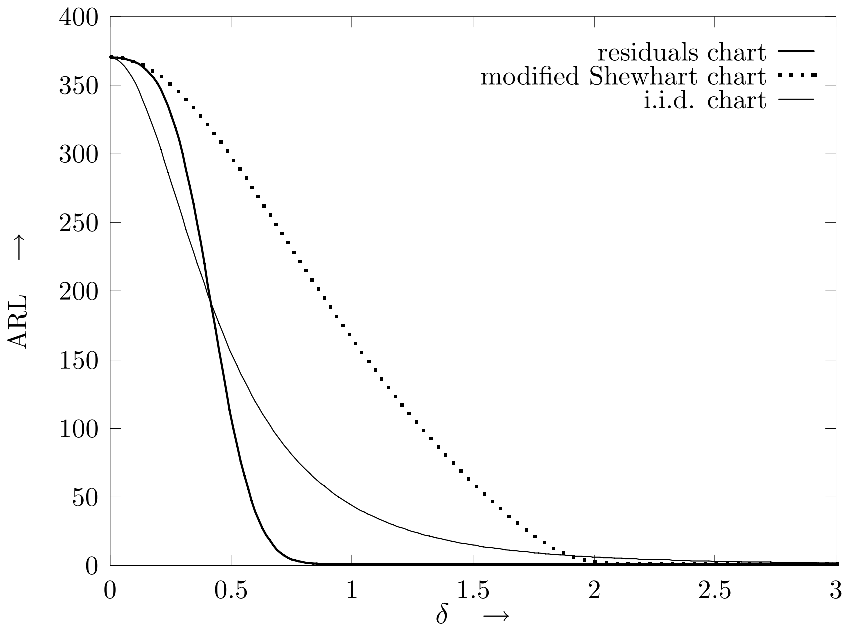

Figure 4.7: Various ARL curves of AR(1) process with $\phi = 0.99$.

ARL. However, for positive ϕ, the modified Shewhart chart is performing better than the residuals control chart.

The findings of the last two sections can be summarized in the advice to use the modified residuals chart for detecting a shift in the mean of AR(1) data with positive ϕ, and to use the residuals chart in case of AR(1) data with negative ϕ. For $\phi < 0$, a shift in the mean is then on average detected faster than in the i.i.d. case, while for $\phi > 0$ not much efficiency is lost relative to the i.i.d. case, provided that ϕ is not too large.

5 A modification of the residuals control chart

In the previous section it was concluded that the bad performance of the residuals chart for positive ϕ is caused by the fact that only a fraction of the shift in the mean is transferred to the residuals. As a result the 'signal to noise' ratio is smaller than δ for $t > T$ and $\phi > 0$ (see formula (4.6)). This has a negative effect on the performance of the residuals control chart. In general, a higher 'signal-to-noise' ratio will result in a more efficient control chart. For example, in Subsection 3.3 we concluded that the efficiency of the well known EWMA chart is mainly due to a good 'signal-to-noise' ratio. The bad 'signal-to-noise' ratio for positive ϕ is an important disadvantage of the residuals chart.

On the other hand, the residuals chart is theoretically very appealing because it takes the serial correlation explicitly into account, and reduces the problem to the well known case of detecting a shift in the mean of independent observations.

5.1 The idea

In this section, we try to combine the theoretical appeal of the residuals chart with a good 'signal-to-noise' ratio. We suggest a modification of the residual chart that roughly maintains independence of the residuals, while the 'signal-to-noise' ratio is approximately δ within a few observations after the shift. In this way, the main drawback of the residuals chart is overcome, and serial correlation is explicitly accounted for. The modified residuals can be monitored using control charts that were designed for detecting a change in the mean of independent observations.

Our suggestion is to plot u_t at time t, where u_t is defined for $t = 0, 1, 2, \cdots$ as

$$u_t \equiv Y_t - \phi Y_{t-1} + \phi \hat{\mu}_t, \tag{5.1}$$

where $\hat{\mu}_t$ is an estimator of μ_t that quickly responds to changes in the mean of the process, such as an exponentially weighted moving average. The rationale behind this suggestion is the following. Suppose that $\hat{\mu}_t$ is a very good estimator for μ_t, so that

$$\hat{\mu}_t \doteq \begin{cases} \mu & \text{for } t = 0, 1, \cdots, T-1 \\ \mu + \delta\sigma_Y & \text{for } t = T, T+1, \cdots. \end{cases} \tag{5.2}$$

In that case

$$u_t \doteq \begin{cases} \mu + \varepsilon_t & \text{for } t = 0, 1, \cdots, T-1 \\ \mu + \delta\sigma_Y + \varepsilon_t & \text{for } t = T, T+1, \cdots. \end{cases} \tag{5.3}$$

The right hand side of (5.3) represents an ideal situation, where the successive u_t are independent, and the shift in μ_t is fully transferred.

Of course, in practice, such a perfect estimator of $\{\mu_t\}$ is not available. Simulation studies of the ARL curve of the modified residuals chart have shown that an exponentially weighted moving average (EWMA, see also Subsection 3.3) is a better option than a regular moving average. How to choose the EWMA smoothing parameter λ will be discussed in Subsection 5.3.

5.2 Comparison to other procedures

To judge the effect of the modification, we compare the ARL curve of the modified residuals chart to ARL curves of the control charts we discussed earlier. In each of the Figures 5.1 through 5.6, ARL curves corresponding to the four different control charts are drawn for a fixed value of ϕ ranging from $\phi = -0.9$ to $\phi = 0.9$. Again, the graph for $\phi = 0$ is left out because all curves coincide.

For all of the curves it is assumed that the the first observation is taken at the time of the shift. The ARL curve for the modified Shewhart chart is computed using formula (4.5), the ARL curve for the residuals chart is $L^*_\phi(\delta, s)$. The ARL curve for the modified residuals chart is derived by simulation. The curve consists of 101 points, which are means of 10,000 replications. The random number generator we used is described in Knypstra (1997). A value of $\lambda = 0.1$ was chosen for computation of the EWMA.

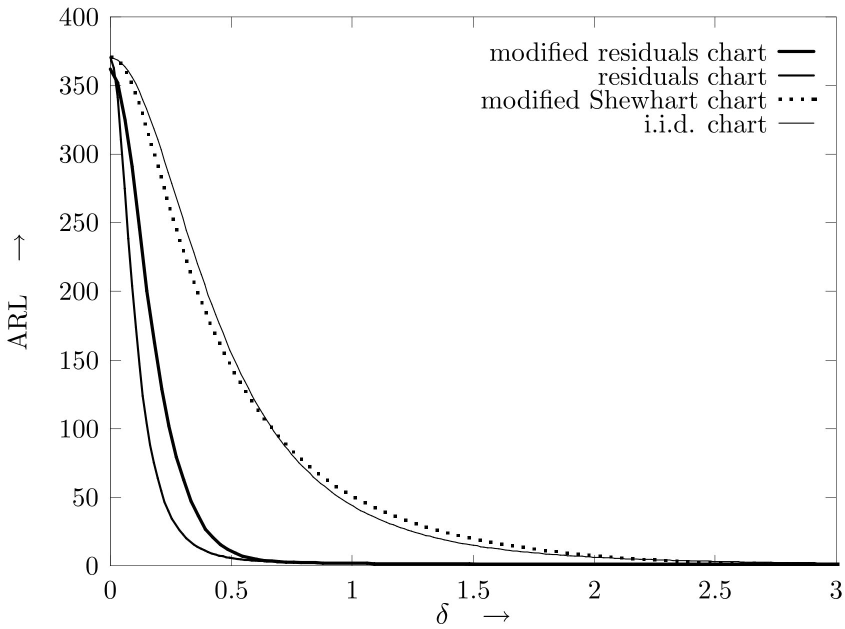

Figure 5.1: Various ARL curves for AR(1) process with $\phi = -0.9$.

From Figures 5.1 through 5.6 we conclude that for negative ϕ, the performance of the modified residuals is better than the modified Shewhart chart and the Shewhart chart for independent observations. However, the excellent performance of the residuals chart for negative ϕ is not equalled by the modified residuals chart. Hence, for $\phi < 0$, the residuals chart remains the best choice, and the modified residuals chart is a well performing second-best

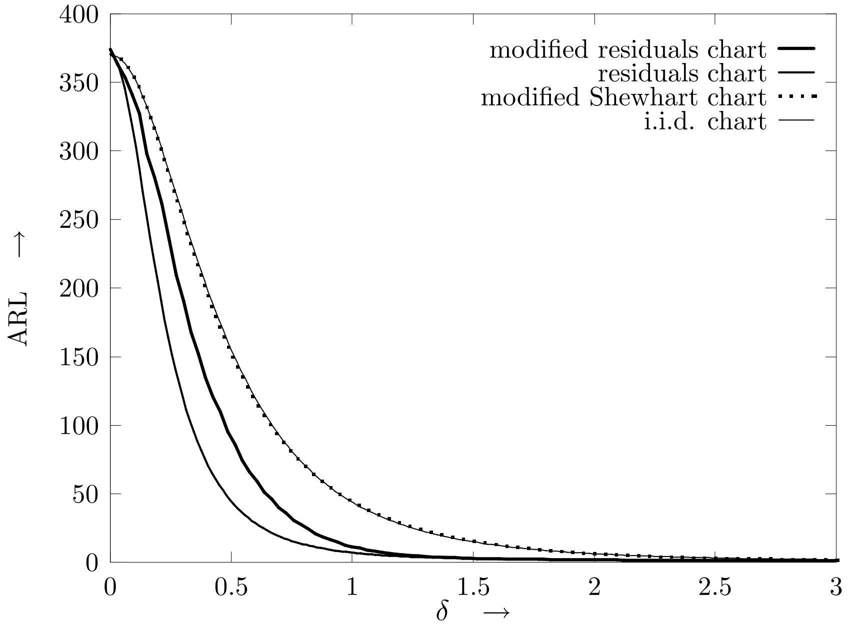

Figure 5.2: Various ARL curves for AR(1) process with $\phi = -0.6$.

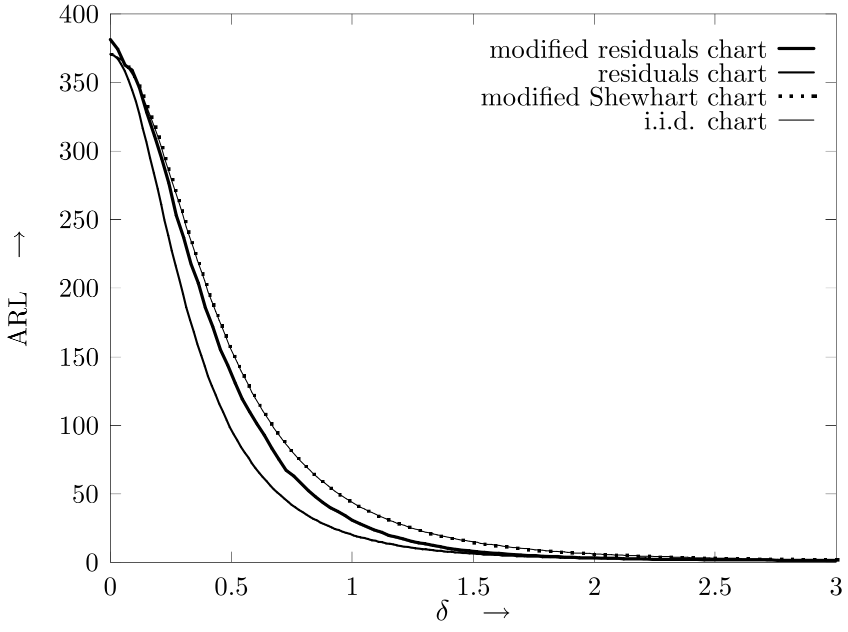

Figure 5.3: Various ARL curves for AR(1) process with $\phi = -0.3$.

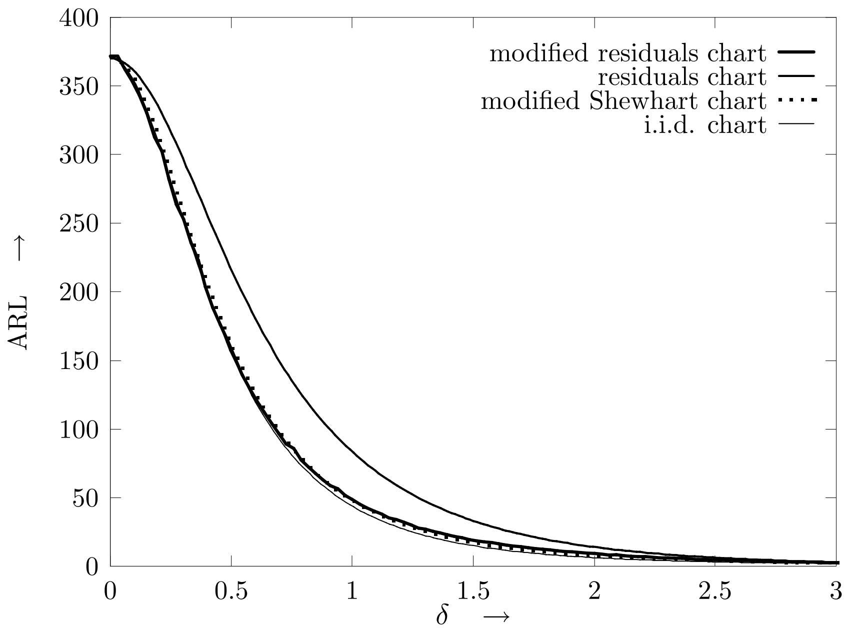

Figure 5.4: Various ARL curves for AR(1) process with $\phi = 0.3$.

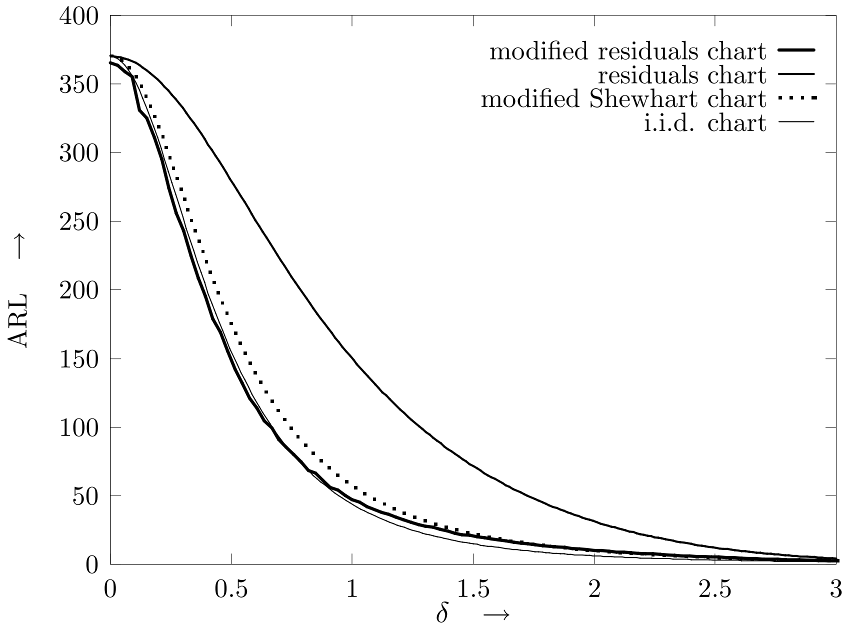

Figure 5.5: Various ARL curves for AR(1) process with $\phi = 0.6$.

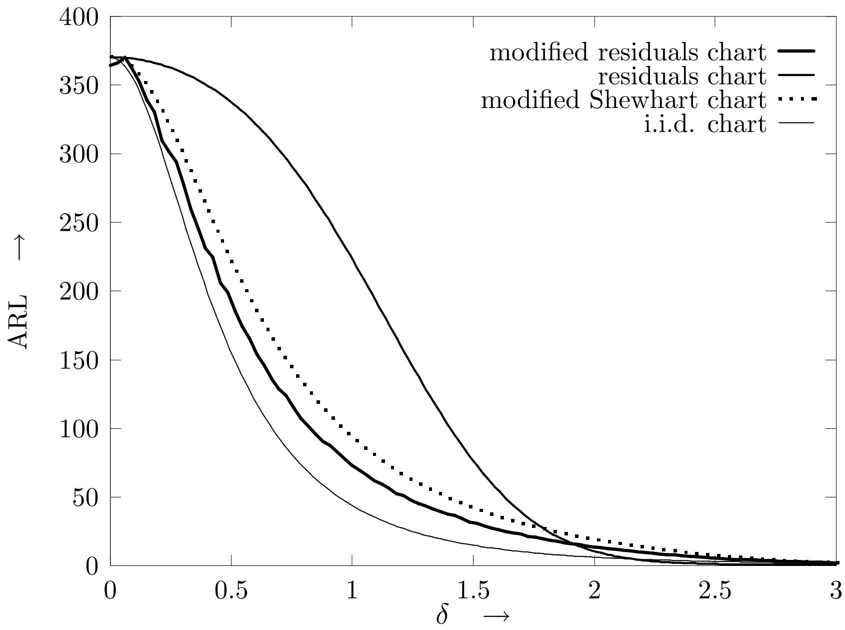

Figure 5.6: Various ARL curves of AR(1) process with $\phi = 0.9$.

option.

For positive ϕ, the modified residuals chart outperforms both the modified Shewhart chart and the residuals chart. The loss of efficiency due to serial correlation relative to the i.i.d. case is negligible for $\phi \geq 0.6$. However, for larger positive values of ϕ this procedure becomes less efficient, too.

5.3 Choice of the EWMA smoothing parameter

For a good performance of the suggested modification, it is necessary to have a good estimator for $E(Y_t)$ available, one that quickly adapts to persisting changes which may occur due to the presence of special causes of variation. In various simulation studies, we experimented with a regular moving average with a small window size of say, 5 to 10 observations. From these simulations we learned that a regular moving average is not the best option: hardly any efficiency is gained for positive ϕ relative to the modified Shewhart chart. Experiments with the EWMA show that this alternative is preferable. However, the choice of the value of λ requires some care. In table 5.1, the effect of the choice of λ on the ARL of the modified residuals chart for an AR(1) process with $\phi = 0.9$ is summarized. This value of ϕ was chosen since for the values of ϕ considered, this ARL curve could use some improvement.

From table 5.1 we conclude that to a limited extend, the choice of λ can be used as a design parameter for the modified residuals chart. If large shifts of size $2\sigma_Y$ or $3\sigma_Y$ are to be detected quickly, a small value of $\lambda = 0.01$ is recommended. A value of $\lambda = 0.05$ is a good choice for detecting a shift of size σ_Y.

Table 5.1: The effect of λ on the ARL of the modified residuals chart for $\phi = 0.9$.

	ARL(0)	ARL(1)	ARL(2)	ARL(3)
$\lambda = 0.01$	377.4 (3.94)	77.3 (0.67)	9.7 (0.22)	1.1 (0.02)
$\lambda = 0.025$	372.1 (3.75)	69.3 (0.62)	10.6 (0.19)	1.8 (0.03)
$\lambda = 0.05$	366.9 (3.76)	67.4 (0.64)	12.3 (0.18)	1.8 (0.04)
$\lambda = 0.075$	368.6 (3.75)	70.6 (0.71)	13.4 (0.18)	2.1 (0.05)
$\lambda = 0.1$	364.4 (3.68)	73.1 (0.74)	13.7 (0.18)	2.3 (0.05)
$\lambda = 0.125$	365.0 (3.63)	75.0 (0.77)	14.2 (0.18)	2.4 (0.05)
$\lambda = 0.15$	364.7 (3.65)	78.0 (0.80)	14.3 (0.19)	2.5 (0.05)

5.4 Discussion

In summary, the modified residuals chart has an overall good performance: if ϕ is negative, it is more efficient than the Shewhart chart for independent observations. It also performs better than the modified Shewhart chart. The efficiency gain of the residuals chart for negative ϕ is only partly attained by the modified residuals chart. For positive ϕ, it outperforms both the residuals chart and the modified Shewhart chart. For small to moderate ϕ there is virtually no loss of efficiency compared to the benchmark curve of the Shewhart chart for independent observations.

The crux of the modification is the addition or subtraction of a portion of $\hat{\mu}_t$. The estimator that is used needs to adapt quickly to persisting changes in the level of the observations, but must be insensitive to the effect of short term random disturbances. As it turns out, an EWMA is a better choice than a regular moving average. The smoothing parameter λ should be chosen somewhere within the range $[0.01, 0.15]$. Within this range, it is to a limited extend possible to choose λ in such a way that the modified residuals chart is most sensitive for shifts of a given size. Table 5.1 can provide some guidance when choosing a value of λ.

In practice, it is of course not possible to estimate the sequence $\{\mu_t\}$ perfectly. If this were possible, formula (5.3) would become an exact relationship, and the corresponding ARL curve would equal the curve of the Shewhart chart for independent observations. Hence, the latter may be viewed as a kind of limiting ARL curve for the modified residuals chart with a very good smoothing procedure. In the simulation studies leading up to this paper, we only considered regular moving averages with a small window size and the EWMA. It is possible that another smoothing procedure performs even better for large positive values of ϕ, in the sense that the corresponding ARL curve is closer to the ARL of the Shewhart chart for independent observations. This remains to be investigated.

For practical purposes, the modified residuals chart is an improvement over existing procedures since the chart outperforms both other Shewhart type control charts for the case of $\phi > 0$. We believe that this case is more likely to occur in practice than the case of negative ϕ. But also from a theoretical point of view this approach is appealing. The extra information on the data structure that is provided by the presence of serial correlation is used explicitly, and the problem is transformed approximately into the more familiar case of monitoring a sequence of independent observations. However, due to the imperfect estimate of $\{\mu_t\}$, some serial correlation remains, and some ad-hoc adjustments to the control limits are needed to attain a desired in-control ARL.

6 Examples

In this section, we illustrate the use of previously discussed control charts by two examples. The first is a real life example, based on a data set that appeared in Shewhart (1931). In his treatment of the data set, Shewhart did not take the presence of serial correlation into account. By using the control charts discussed in this paper, we arrive at other conclusions than Shewhart did. The second example is a based on a simulated AR(1) series with a persistent change in the mean of the observations.

6.1 A real life example

The first book on quality control stems from the year 1931. It is written by the developer of the control chart: Dr. Walter A. Shewhart. In this very well written work, the newly developed concepts of quality control are illustrated with real-life examples. The second data set that appears in this book consists of 204 observations of the resistance of a certain insulation material. In table 6.1, these observations are reprinted. The observations

Table 6.1: Electrical resistance of insulation in megohms

5,045	4,635	4,700	4,650	4,640	3,940	4,570	4,560	4,450	4,500	5,075	4,500
4,350	5,100	4,600	4,170	4,335	3,700	4,570	3,075	4,450	4,770	4,925	4,850
4,350	5,450	4,110	4,255	5,000	3,650	4,855	2,965	4,850	5,150	5,075	4,930
3,975	4,635	4,410	4,170	4,615	4,445	4,160	4,080	4,450	4,850	4,925	4,700
4,290	4,720	4,180	4,375	4,215	4,000	4,325	4,080	3,635	4,700	5,250	4,890
4,430	4,810	4,790	4,175	4,275	4,845	4,125	4,425	3,635	5,000	4,915	4,625
4,485	4,565	4,790	4,550	4,275	5,000	4,100	4,300	3,635	5,000	5,600	4,425
4,285	4,410	4,340	4,450	5,000	4,560	4,340	4,430	3,900	5,000	5,075	4,135
3,980	4,065	4,895	2,855	4,615	4,700	4,575	4,840	4,340	4,700	4,450	4,190
3,925	4,565	5,750	2,920	4,735	4,310	3,875	4,840	4,340	4,500	4,215	4,080
3,645	5,190	4,740	4,375	4,215	4,310	4,050	4,310	3,665	4,840	4,325	3,690
3,760	4,725	5,000	4,375	4,700	5,000	4,050	4,185	3,775	5,075	4,665	5,050
3,300	4,640	4,895	4,355	4,700	4,575	4,685	4,570	5,000	5,000	4,615	4,625
3,685	4,640	4,255	4,090	4,700	4,700	4,685	4,700	4,850	4,770	4,615	5,150
3,463	4,895	4,170	5,000	4,700	4,430	4,430	4,440	4,775	4,570	4,500	5,250
5,200	4,790	3,850	4,335	4,095	4,850	4,300	4,850	4,500	4,925	4,765	5,000
5,100	4,845	4,445	5,000	4,095	4,850	4,690	4,125	4,770	4,775	4,500	5,000

Source: Shewhart (1931), page 20, Table 2.

are taken in the order in which they are observed. Shewhart decides to take subgroups of size four, and presents a control chart for the mean. The 51 subgroup averages are compared to control limits ‘within which experience has shown that these observations should fall’. Since this data set is presented in one of the introductorily chapters, Shewhart does not explain precisely how the control limits are computed. However, on page 296, Shewhart suggests computing the control limits as

$$\bar{\bar{X}} \pm 3\frac{\hat{\sigma}}{\sqrt{n}},$$

which is a formula that is very familiar to most SPC practitioners. The central line is the overall mean, which can be computed as 4,498MΩ. Estimating σ as the mean of the sample standard deviations of the subgroups, corrected by the constant $c_4(4)$, results in a lower control limit of 4.006MΩ, and an upper control limits of 4,991MΩ. These values agree closely with the control limits that Shewhart depicted in the corresponding control chart. Eight of the subgroup averages fall outside the control limits, see Figure 6.2(a). Shewhart interprets these out-of-control signals as ‘an indication of the existence of causes of variability which could be found and eliminated'. He reports that further research was instituted to find these causes of variability. The search was successful and a second control chart is presented, wherein data points are depicted that were taken after elimination of these causes. All values remain within much tighter limits, and Shewhart concludes that 'this variation should be left to chance'.

However, if we take a closer look at the data set in table 6.1, it appears that the successive values exhibit serial correlation. In fact, the data set appears to be a typical example of observations that can be successfully modelled using an AR(1) model. The sample autocorrelation function is exponentially declining, and the sample partial autocorrelation function shows a single spike at lag 1, see Figure 6.1.

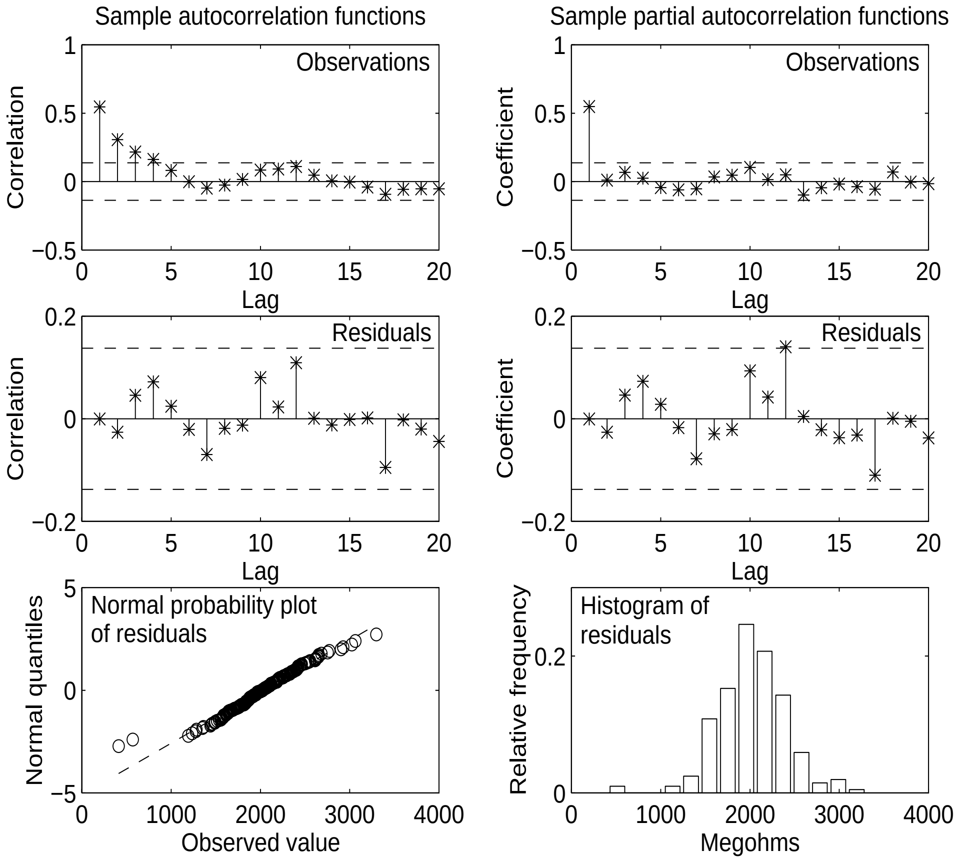

Figure 6.1: Analysis of the data in table 6.1.

The AR parameter ϕ can be estimated as $\hat{\phi} = 0.549$. In Figure 6.1 it is shown that the residuals of this model show no significant serial correlation.

A normal probability plot of the residuals indicates the presence of two or three outliers. All other observations lie more or less on a straight line, even after removal of the suspected data points. The histogram of the residuals is a little skewed to the right. Nevertheless, we feel it is safe to assume that the residuals of this model are uncorrelated and normally distributed.

The autocorrelation function of the 51 subgroup averages shows less convincingly that serial correlation is present. Here we observe a well known phenomenon: by taking subgroup averages, the serial correlation is reduced. Ignoring the serial correlation in the data seems therefore justifiable. But, as we will see, by taking the serial correlation into account, less out-of-control signals are generated. Wardell, Moskowitz, and Plante (1992) warn for this kind of inconsiderate subgroup taking:

> "*However, if the data are truly autocorrelated, the points on the Shewhart chart will still show runs which are essentially due to correlation resulting from common causes rather than any special cause*".

In Figure 6.2, four control charts corresponding to the data in table 6.1 are depicted. Figure 6.2(a) shows a control chart of subgroup averages, as proposed by Shewhart (1931). In Figures 6.2(b)–6.2(d), the modified Shewhart control chart, the residuals control chart and the modified residuals control chart are depicted, respectively.

The control limits in Figures 6.2(b)–6.2(d) are adjusted to have an approximate in-control ARL of 370.4. This was presumably Shewhart's intention. In Figure 6.2(b), two sets of control limits are depicted. The dashed limits are computed using an residuals-based estimate of σ_Y, which is in our opinion preferable to an estimate based on the correlated observations. The dotted control limits correspond to the latter estimate.

The reason why we prefer estimating σ_Y on the basis of residuals and not on the basis of the correlated observations is the following. From the results in Anderson (1971), it can be shown that the expectation of

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (Y_i - \bar{Y})^2$$

where $Y_1, \cdots, Y_n$ is a sequence of observations of a stationary AR(1) process, is equal to

$$\mathrm{E}(S^2) = \sigma_Y^2 - \frac{2}{n-1} \sum_{r=1}^{n-1} \left(1 - \frac{r}{n}\right) \frac{\phi^r}{\sigma_Y^2}.$$

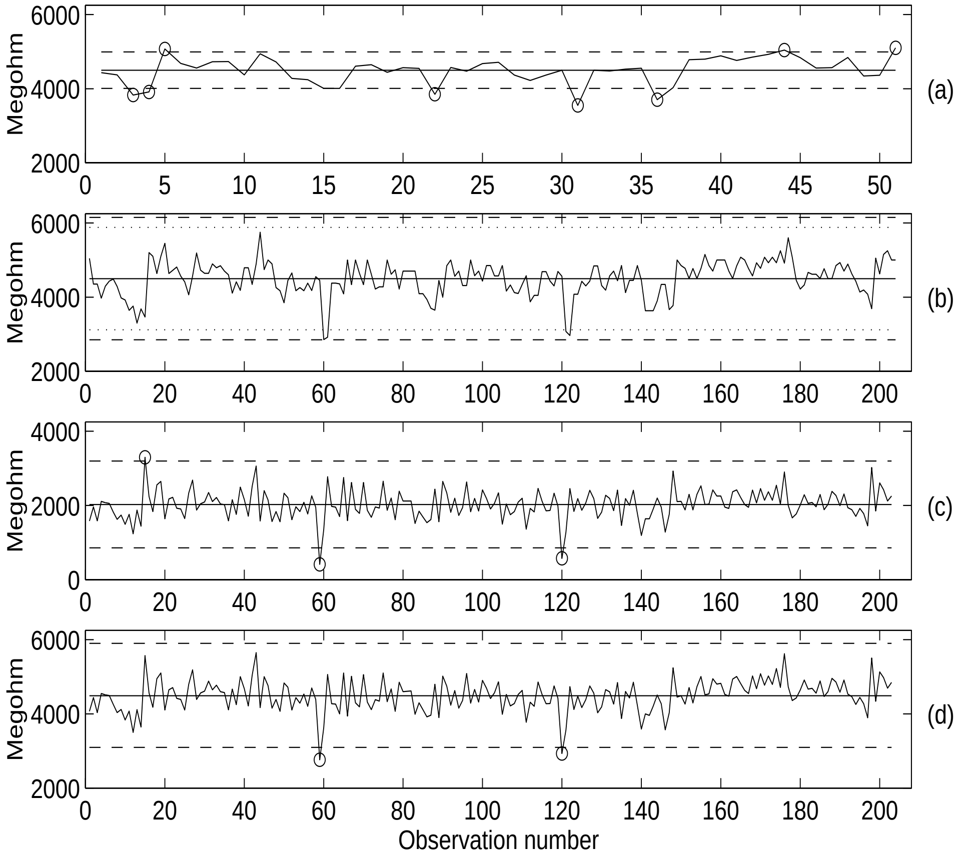

Figure 6.2: Four control charts corresponding to table 6.1. (a)=Chart for subgroup averages, (b)=modified Shewhart, (c)=residuals chart, (d)=modified residuals chart.

For positive ϕ, this means that $\mathrm{E}(S^2)$ is biased downwardly. For $n \rightarrow \infty$ the bias disappears, since the process is assumed to be stationary. The downward bias of $\mathrm{E}(S^2)$ explains why the control limits in Figure 6.2(a) are so much tighter than the control limits in the other control charts. The control limits in Figure 6.2(a) are computed using estimates of σ_Y based on $n = 4$ correlated observations. For the residual control charts in Figure 6.2(c) and 6.2(d), the standard deviations of approximately uncorrelated residuals determine the width of the control limits.

The downward bias in $\mathrm{E}(S^2)$ also explains the difference between the two sets of control limits in Figure 6.2(b). However, since $n = 204$ is large, the difference is small in this case.

The control charts in Figures 6.2(c) and 6.2(d) are residual control charts. In the previous sections, we argued that these charts have a high probability of detecting a shift at the first observation. This agrees with the control charts in Figure 6.2, where both Figures 6.2(c) and 6.2(d) generate an out-of-control signal, while the modified Shewhart chart does not.

Based on the results in this subsection, we have reason to suspect observations 60, and 121 and perhaps observation 16, too. The out-of-control signals on the last three control chart in Figure 6.2 indicate the presence of special causes of variation, each causing a single spike in the mean of the process. In this paper, we concentrated mainly on persistent changes in the mean of the process. The example discussed in this section is nevertheless useful since it clearly shows the pitfalls of ignoring the presence of serial correlation.

6.2 A simulated example

In this subsection, we will illustrate the use of the control charts that were considered in this paper on a simulated AR(1) sequence with a persistent shift in the mean. To this end, we simulated 150 observations of an AR(1) process with $\phi = 0.6$. The expectation of the first 79 observations is 2, at observation 80, a shift to $\mu = 3.25$ is introduced, which corresponds to a shift in the mean of the process of $1\sigma_Y$. In Figure 6.3, respectively the modified Shewhart chart, the residuals chart, and the modified residuals chart are used to monitor these observations.

Figure 6.3(a) shows that the modified Shewhart chart signals at observation 93. The residuals chart in Figure 6.3(b) signals at the 84th and the 92nd residual. For the modified residuals chart in Figure 6.3(c) a value of $\lambda = 0.1$ is used for the EWMA. Out-of-control signals are observed at 84th, the 92nd, 100th, and the 122nd residual. The control limits are in all cases chosen such that the in-control ARL is approximately 370.

The example shows that for this sequence of observations, the modified residuals chart signals the shift in the mean early, and generates the largest

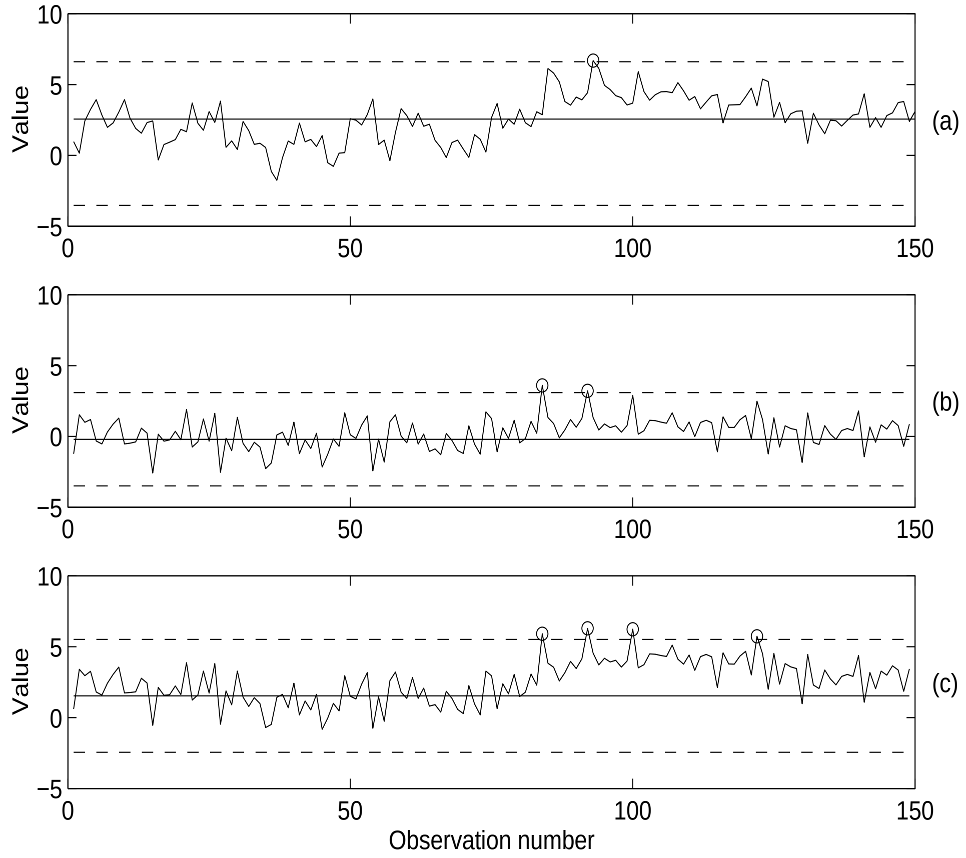

Figure 6.3: Control charts for a simulated AR(1) series.
(a)=modified Shewhart chart,
(b)=residuals chart, (c)=modified residuals chart.

number of out-of-control signals.

7 Conclusions

In this paper we discussed three Shewhart type control charts that are able to take first order autocorrelation in the observations into account. The performance of these control charts was compared by means of ARL curves.

The first control chart that was evaluated is a modification of the classical Shewhart control chart. It basically ignores the autocorrelation; the control limits are in an rather ad-hoc manner adjusted to attain a certain in-control ARL. For negative and moderate positive first order autocorrelation, the ARL performance of this control chart is comparable to that of the classical Shewhart control chart for independent observations. This is a comforting observation: out-of-control signals of the modified Shewhart chart can be interpreted just as in the i.i.d. case in a large number of practical situations. However, the fact that the extra information on the data structure remains unused is a little unsatisfactory.

Secondly, the residuals chart was discussed. This chart utilizes the residuals of a fitted time series model to monitor the process for shifts in the mean. If the time series model is appropriate for the data, the residuals are approximately uncorrelated. In this way, the serial correlation is explicitly taken into account, and the problem is reduced to the well known case of detecting a shift in the mean of independent observations. For negative first order autocorrelation and for values of ϕ that are extremely close to one, the performance of the residuals chart is excellent. Compared to the case of independent observations, a shift in the mean is on average detected faster. However, for values of $\phi > 0$ and not close to one, the residuals chart behaves poorly. This is an important drawback of the residuals chart.

To overcome this drawback, a third control chart is introduced in this paper. It is a modification of the residuals chart. This chart was shown to have a good overall ARL performance. For negative values of ϕ it is not as efficient as the regular residuals chart, but it is more efficient than the modified Shewhart chart and the classical Shewhart chart for the i.i.d. case. For positive ϕ, it outperforms both the modified Shewhart chart and the residuals chart. The difference with the classical Shewhart chart for independent observations is negligible for a large region of ϕ-values. For large ϕ (say $\phi > 0.8$) however, improvement of ARL performance remains desirable. Further research into this is needed.

Examples were discussed to illustrate the use of these control charts. The first example taken from Shewhart (1931), showed that ignoring serial correlation in the data may result in misleading conclusions. The example underlines the importance of developing control charts that are able to take serial correlation into account.

In the second example, the three control charts were used to monitor a simulated AR(1) sequence for a shift in the mean. The results agreed with the behavior that was expected on the basis of the ARL comparisons.

Finally, we would like to make two general remarks. Firstly, in this paper, only Shewhart-type control charts were discussed. For each of the procedures considered, improvement of the ARL performance is possible by utilizing CUSUM or EWMA control schemes instead. Some work concerning comparison of the ARL performance of such control schemes has been carried out by other authors. However, this remains a topic for further research. Secondly, throughout this paper we assumed that the process parameters were known. In practice, these have to be estimated. In Kramer and Schmid (1996) it is shown that both the modified Shewhart and the residuals chart react sensibly to parameter estimation. The robustness for parameter estimation of the modified residuals chart needs to be investigated.

References

Alwan, L. C. and H. V. Roberts (1988), "Time-series modeling for statistical process control", *Journal of Business & Economic Statistics*, **6**(1), 87–95.

Anderson, T. W. (1971), *The Statistical Analysis of Time Series*, John Wiley & Sons, New York.

Bagshaw, M. and R. A. Johnson (1975), "The effect of serial correlation on the performance of CUSUM tests II", *Technometrics*, **17**(1), 73–80.

Berthouex, P. M., W. G. Hunter, and L. Pallesen (1978), "Monitoring sewage treatment plants: Some quality control aspects", *Journal of Quality Technology*, **10**(4), 139–149.

Brook, D. and D. A. Evans (1972), "An approach to the probability distribution of CUSUM run length", *Biometrika*, **59**(3), 539–549.

Crowder, S. V. (1987), "A simple method for studying run-length distributions of exponentially weighted moving average charts", *Technometrics*, **29**(4), 401–407.

Harris, T. J. and W. H. Ross (1991), "Statistical process control procedures for correlated observations", *The Canadian Journal of Chemical Engineering*, **69**, 48–57.

Harvey, A. C. (1993), *Time Series Models*, second edition, Harvester Wheatsheaf, New York.

Hunter, S. J. (1986), "The exponentially weighted moving average", *Journal of Quality Technology*, **18**(4), 203–210.

Johnson, R. A. and M. Bagshaw (1974), "The effect of serial correlation on the performance of CUSUM tests", *Technometrics*, **16**(1), 103–112.

Knypstra, S. (1997), "Kansverdelingen en objecten", *Kwantitatieve Methoden*, **18**(54), 91–95.

Kramer, H. and W. Schmid (1996), "Control charts for time series", Arbeitsbericht 59, Europa-Universität Viadrina Frankfurt (Oder), Fakultät für Wirtschaftswissenschaften.

Kramer, H. and W. Schmid (1996), "The influence of parameter estimation on the ARL of Shewhart type charts for time series", Arbeitsbericht 60, Europa-Universität Viadrina Frankfurt (Oder), Fakultät für Wirtschaftswissenschaften.

Lucas, J. M. and M. S. Saccucci (1990), "Exponentially weighted moving average control schemes: Properties and enhancements", *Technometrics*, **32**(1), 1–29.

Montgomery, D. C. (1996), *Introduction to Statistical Quality Control*, third edition, John Wiley & Sons, New York.

Montgomery, D. C. and C. M. Mastrangelo (1991), "Some statistical process control methods for autocorrelated data", *Journal of Quality Technology*, **23**(3), 179–204.

Roberts, S. W. (1959), "Control chart tests based on geometric moving averages", *Technometrics*, **1**(3), 239–250.

Ryan, T. P. (1989), *Statistical Methods for Quality Improvement*, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, New York.

Ryan, T. P. (1991), "Discussion of *some statistical process control methods for autocorrelated data*", *Journal of Quality Technology*, **23**(3), 200–202.

Schmid, W. (1995), "On the run length of a Shewhart chart for correlated data", *Statistical Papers*, **36**, 111–130.

Schmid, W. and A. Schöne (1997), "Some properties of the EWMA control chart in the presence of autocorrelation", *The Annals of Statistics*, **25**(3), 1277–1283.

Shewhart, W. A. (1931), *Economic Control of Quality of Manufactured Product*, D. van Nostrand Company, Inc., Toronto.

Vasilopoulos, A. V. and A. P. Stamboulis (1978), "Modification of control chart limits in the presence of data correlation", *Journal of Quality Technology*, **10**(1), 20–30.

Wardell, D. G., H. Moskowitz, and R. D. Plante (1992), "Control charts in the presence of data correlation", *Management Science*, **38**(8), 1084–1105.