COMPUTATION OF THE PERFORMANCE OF SHEWHART CONTROL CHARTS

Pieter Mulder, Julian Morris and Elaine B. Martin

Centre for Process Analytics and Control Technology,
School of Chemical Engineering and Advanced Materials,
University of Newcastle, Newcastle upon Tyne, NE1 7RU, UK

Abstract: The performance of a control chart in statistical process control is often quantified in terms of the Average Run Length (ARL). The ARL enables a comparison to be undertaken between various monitoring strategies. These are often determined through Monte Carlo simulation studies. Monte Carlo simulations are time consuming and if too few runs are performed then the results will be inaccurate. An alternative approach is proposed based on analytical computation. The analytical results are compared with those of the Monte Carlo simulations for three case studies.

Keywords: Statistical Process Control, Control charts, Serial correlation, Monte Carlo simulations

1. INTRODUCTION

In Statistical Process Control (SPC), a variety of control charts have been applied including Shewhart, CUSUM and EWMA (e.g. Montgomery, 1991; Wetherill and Brown, 1991). Each method has associated advantages and disadvantages that have been reported in the literature. Control chart performance is traditionally quantified in terms of the Average Run Length (ARL). Run length is defined as the number of observations from the start of the control chart to the first out-of-control signal.

Except for simple cases (e.g. Brook and Evans, 1972; Schmid, 1995), Monte Carlo simulations have been used to determine the ARL. This involves the realisation of a vector containing a random signal and then applying the control scheme and measuring the run length (time to the first alarm). This is repeated many times, each time a different random vector is generated, and finally the ARL is computed. The main issue with this method is the trade-off between computer time and accuracy of the results. A large number of realisations are necessary if the results are to be precise (Lowry *et al*, 1992; Wardell *et al*, 1994). Therefore it is believed that an analytical method for the computation of the ARL would be desirable. The density function of the run length of a control chart is first constructed based on the in-control probability of an observation. This approach is similar to that of Wetherill and Brown (1991). They assumed that the in-control probability was constant for every observation. In contrast, in this work, this constraint is relaxed. This allows the computation of the ARL for more complicated SPC monitoring strategies, and ultimately for correlated data.

The ARL is investigated in more detail for three case studies. The first case study looks at the ARL of a Shewhart control chart based on independent data and is derived for both the in-control and out-of-control situation. This example demonstrates the validity of the approach. The impact of serial correlation on the performance of control charts is well known (Alwan and Roberts, 1988; Montgomery and Mastralango, 1991). One solution is to estimate an ARMA model (Harris and Ross, 1991) for univariate systems, or a VAR model (Mulder *et al*, 2001) for multivariate systems, and to monitor the residuals, which are free of serial correlation. In the second case study, the focus is on the residuals of a first order AutoRegressive, AR(1), time series model as defined by Box *et al,* (1994).

For the third case study, the ARL of a correlated time series generated by an AR(1) model is computed. Schmid (1995) claimed that an explicit solution does not exist for the ARL of correlated data, and that only general statements about the ARL are possible. In this study it will be shown that although there is not an explicit solution for the ARL, there is a numerical approximation. The analytical results are compared with the results of Monte Carlo simulations for each of the three case studies.

2. CONTROL CHARTS

The run length of a control chart is defined as the number of observations until the first observation moves outside of the control limits. After this observation, the control chart is stopped and calculation of the run length is recommenced from the next in-control observation. In this section, the density function of the run length is constructed.

The probability that an observation, X_k, is in control at time point, *k*, is given by:

$$P(UCL > X_k > LCL) \tag{1}$$

and the probability that at point, *k*, observation, X_k, is out-of-control is defined as:

$$\begin{aligned} &P(X_k > UCL \; or \; X_k < LCL) \\ &=1- P(UCL > X_k > LCL) \end{aligned} \tag{2}$$

Also it is assumed that an observation is either in-control or out-of-control. In a control chart, an observation is only recorded if the previous point was in-control. That is an observation can only be deemed to be in-control at time point *k* if the observations at $1,2,\ldots,k-1$ were in control:

$$\begin{aligned} P_{IC,k} &= P(UCL > X_1 > LCL)\cdot P(UCL > X_2 > LCL)\cdot \ldots \\ &\cdot P(UCL > X_{k-1} > LCL)\cdot P(UCL > X_k > LCL) \end{aligned} \tag{3}$$

Also an observation at time point *k* in a control chart is out-of-control if:

$$\begin{aligned} P_{OC,k} &= P(UCL > X_1 > LCL)\cdot P(UCL > X_2 > LCL)\cdot \ldots \\ &\cdot P(UCL > X_{k-1} > LCL)\cdot P(X_k > UCL \; or \; X_k < LCL) \end{aligned} \tag{4}$$

Based on equations 3 and 4, the Average Run Length is the expectation of the out-of-control run length and is given by:

$$E(kP_{OC,k}) = \sum_{k=1}^{\infty} kP_{IC,k-1}P(X_k > UCL \; or \; X_k < LCL) \tag{5}$$

Based on the following definition

$$P(UCL > X_k > LCL) = \beta_k \tag{6}$$

Equation 2 is given as:

$$P(X_k > UCL \; or \; X_k < LCL) = 1-\beta_k \tag{7}$$

and equation 5 is redefined as:

$$ARL = E(kP_{OC,k}) = \sum_{k=1}^{\infty} k(1-\beta_k)\prod_{j=1}^{k-1}\beta_j \tag{8}$$

This is as described by Wetherill and Brown (1991), except that β in equation 8 can differ for each time point, *k*.

2.1 Case 1 - Independent Data

For independent data, the value of an observation is independent of its previous value, thus $\beta_k = \beta$ for $1,2,\ldots,k-1$, and equation 8 becomes:

$$ARL = \sum_{k=1}^{\infty} k(1-\beta)\beta^{k-1} = \frac{1}{1-\beta} \tag{9}$$

This result agrees with that of Wetherill and Brown (1991).

2.2 Case 2 - Residuals from an AR(1) Model

An AR(1) time series model is given by:

$$\begin{aligned} y_t &= \xi_t + \eta_t \\ \xi_t &= \alpha\xi_{t-1} + e_t \end{aligned} \tag{10}$$

where y_t is the observed data, ξ_t is the underlying correlated time series, with α as its autoregressive parameter, and e_t is a white noise vector with variance σ_e^2, which is assumed to have a Normal distribution, and η_t is the mean shift applied to the data vector y_t (Kaskavelis, 2000). The one-step ahead prediction errors of an AR(1) model are:

$$\begin{aligned} \hat{e}_t &= y_t - y_{t|t-1} \\ &= \alpha\xi_{t-1} + e_t + \eta_t - \alpha\xi_{t-1} - \alpha\eta_t \\ &= e_t + \eta_t - \alpha\eta_{t-1} \end{aligned} \tag{11}$$

When a process is in-control, $\eta_t = 0$, for all *t*, then the probability that at time point *k*, the process is in-control is constant, $\beta_k = \beta$, $k \geq 1$. Therefore the in-control ARL is given by:

$$ARL = \sum_{k=1}^{\infty} k(1-\beta)\beta^{k-1} = \frac{1}{1-\beta} \qquad (12)$$

From equation 12 and for the desired in-control ARL, the control limits for the Shewhart control charts can be derived. For the out-of-control case, it is assumed that a constant mean shift is applied to y_t, $\eta_t = \eta$ for all $k \geq 1$. Thus according to equation 11, at $k = 1$, the probability that, when a mean shift is applied to y_t, the process is in-control is β_1 and for $k \geq 2$, $\beta_k = \beta$. The mean shift in the residuals for $k = 1$ is different to that for $k \geq 2$. Apley and Shi (1999) termed this the fault signature of a step change in the residuals for univariate systems. This issue was not considered by Harris and Ross (1991) or Kaskavelis (2000). They assumed that the mean shift in the residuals is identical for all k. The probability that the control chart will give an out-of-control signal at some point in the future is:

$$\begin{aligned} P &= (1-\beta_1) + \sum_{k=2}^{\infty} \beta_1 (1-\beta)\beta^{k-2} \\ &= (1-\beta_1) + \beta_1(1-\beta)\sum_{k=0}^{\infty}\beta^k = 1 \end{aligned} \qquad (13)$$

since $\sum_k^{\infty} \beta^k = (1-\beta)^{-1}$ for $0 \leq \beta < 1$. The out-of control ARL is given by:

$$\begin{aligned} ARL &= (1-\beta_1) + \sum_{k=2}^{\infty} k\beta_1(1-\beta)\beta^{k-2} \\ &= (1-\beta_1) + \frac{\beta_1(1-\beta)}{\beta}\left(\frac{1}{(1-\beta)^2} - 1\right) \end{aligned} \qquad (14)$$

When $\beta_1 = \beta$, it can be seen that equation 14 is equivalent to equation 9 and 12.

2.3 *Case 3 - Serially Correlated Data*

In this section it will be shown how β_k can be computed for AR(1) processes via the probability distribution function. It is assumed that observations X_k are monitored using a Shewhart control chart with an Upper Control Limit (UCL) and a Lower Control Limit (LCL). The cumulative distribution function of observation X_k at time point k is defined as (Papoulis, 1991):

$$F_k(x) = \int_{-\infty}^{x} f_k(z)dz \qquad (15)$$

where f_k is the probability density function of observation X_k at time point k. The time series structure is defined as in equation 10, where α has variance $1-\alpha^2$. A condition of the control chart is that an observation at k is only plotted if the observation at $k-1$ is in-control:

$$UCL > X_{k-1} > LCL \qquad (16)$$

otherwise the control chart would have been terminated at $k-1$. The conditional distribution function of X at $k-1$ is given by:

$$f_{k-1,cond}(x) = \frac{f_{k-1}(x)}{F_{k-1,uncond}(UCL) - F_{k-1,uncond}(LCL)} \qquad (17)$$

The probability distribution function of the term αX_{k-1} is defined as (Papoulis, 1991):

$$g_k(x) = \frac{1}{|\alpha|} \cdot f_{k-1,cond}(x) \qquad (18)$$

The white noise term e_t is normally distributed with variance $1-\alpha^2$ so that the unconditional X has unit variance. The probability distribution function of e_t is denoted as $N_k(x)$. Since αX_{k-1} and e_t are independent, the probability distribution function of their sum, the probability distribution function of X_k, is given by the convolution product (Papoulis, 1991):

$$f_k(x) = \int_{-\infty}^{x} g_k(z) n_k(z-x)dz \qquad (19)$$

The in-control probability is thus the probability that observation X_k lies between the control limits:

$$\beta_k = F_k(UCL) - F_k(LCL) = \int_{LCL}^{UCL} f_k(z)dz \qquad (20)$$

At the start of a control chart, no other observations are known. Therefore X_1 can be regarded as the unconditional observation of X, and subsequently β_1 is computed from equation 15. For $k > 1$, β_k can be computed recursively by the procedure described above, equations 17 to 19.

3. RESULTS

In this section, the theoretical relationships derived in the previous section are compared with the results from Monte Carlo simulations. For all Monte Carlo simulations 10,000 realisations of the control charts were computed. Each realisation comprised 10,000 observations and the first observation outside the control limits was taken to define the run length for that realisation. Since the ARL is the mean value of the run lengths of the realised control charts, the standard error of the ARL is:

$$\sigma_{ARL} = \frac{ARL}{\sqrt{N}} \tag{21}$$

where N is the number of realisations. Error bars will thus indicate the standard error of the Monte Carlo simulations. For the three cases it is assumed that the metric, the data in case 1 and 3, and the residuals in case 2, are monitored using a Shewhart control chart. For each case it is assumed that the metric used in the control chart is normally distributed and that the desired in-control ARL is equal to 370. This corresponds to a Shewhart control chart with control limits at -3σ and $+3\sigma$.

3.1 Case 1 - Independent Data

The observations, X, are drawn from a population with a normal distribution that are offset by a mean shift, η. The mean of the distribution is equal to η and its variance is σ_X^2:

$$X \sim N(\eta, \sigma_X^2) \tag{22}$$

The probability that X lies between the control limits is given by:

$$\beta = F(UCL) - F(LCL) = \int_{LCL}^{UCL} f(\xi) d\xi \tag{23}$$

where f is the probability distribution function of X, the normal distribution. Subsequently the ARL can be computed from equation 15.

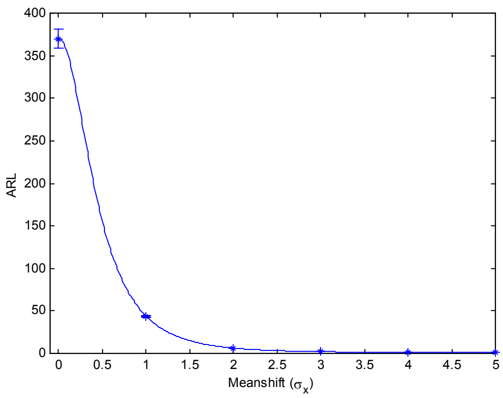

Fig. 1. Performance of Shewhart control chart for independent data.

The ARL as a function of the mean shift is shown in Fig. 1. The dots represent the results of the Monte Carlo simulation with error bars that indicate –3/+3 standard error of the mean and the solid line is the theoretical computation. It can be seen that the theoretical results correspond with those from the Monte Carlo simulations.

3.2 Case 2 - Residuals from an AR(1) Model.

The desired in-control ARL is 370. It is assumed that a step function of size η is superimposed on the time series y_t:

$$\begin{aligned} \eta_t &= 0 \quad t \le 0 \\ \eta_t &= \eta \quad t > 0 \end{aligned} \tag{24}$$

From equation 24, for $k = 1$, $\hat{e}_k \sim N(\eta, \sigma_e^2)$ and for $k \ge 1$, $\hat{e}_k \sim N(\eta(1-\alpha), \sigma_e^2)$. Together with Equation 15, this allows the computation of β for all k. The ARL is computed from equation 14.

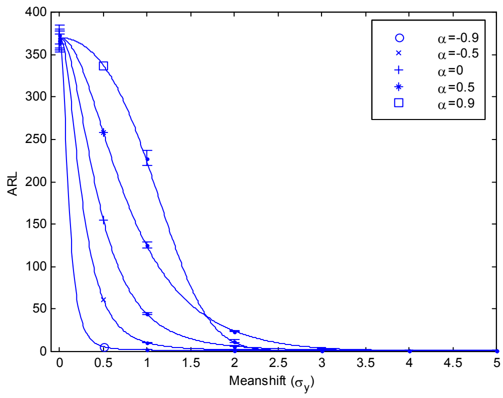

Fig. 2. Performance of Shewhart control chart for the residuals of an AR(1) model.

The ARL as a function of the mean shift for several values of alpha is shown in Fig 2, together with the results from the Monte Carlo simulations. The error bars indicate the results of the Monte Carlo simulations with –3/+3 standard error of the mean. The solid line indicates the theoretical calculation. Again it can be seen that the theoretical results correspond to those from the Monte Carlo simulations.

3.3 Case 3 - Serially Correlated Data.

The ARL as a function of the mean shift was determined in the previous two cases. In this case study, the influence of serial correlation on the in-control ARL was investigated. In contrast to cases 1 and 2, there is no direct analytical relationship for the ARL. Therefore a numerical approach was used. It is assumed that the probability distribution function of X for the first observation is that of the normal distribution with mean zero and unit variance. Based on equations 17 to 20, the probability distribution function of the second observation is computed. Then β_2 is computed from equation 15. These steps are repeated until β_k converges. The values of β_k

are then used in equation 8 to compute the ARL.

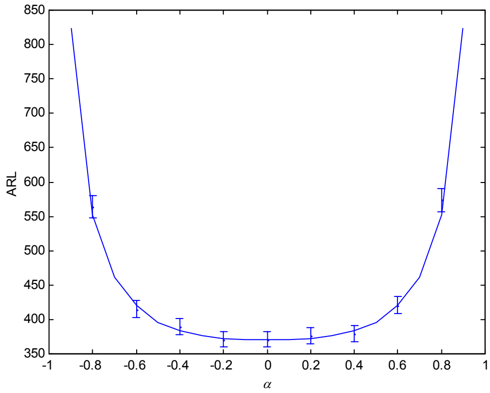

Fig. 3. The in-control ARL of Shewhart control charts for serial correlated data.

The ARL as a function of α is given in Fig. 3, together with the results from the Monte Carlo simulations. The dots represent the Monte Carlo simulations and the error bars indicate –3/+3 standard error of the mean and the solid line indicates the theoretical calculation. Again it can be seen that the theoretical results correspond to those from the Monte Carlo simulations.

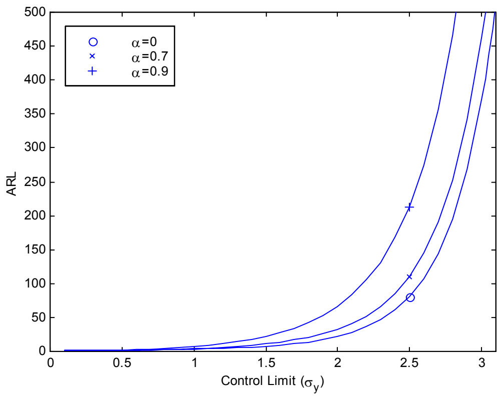

Fig. 4. The in-control ARL as function of the control limit for a selection of values for α.

As can be observed in Fig. 3, the in-control ARL only depends on the absolute value of α. It also can be seen that for increasing α, the in-control ARL increases. This means that on average it takes longer to detect a false alarm. Although this might appear advantageous, in practice the in-control ARL can be considered as a design parameter since it is implied by the choice of significance level δ. For independent and identical data, the in-control ARL is 1/δ, but for serially correlated data this relationship is not valid. Kaskavelis (2000) proposed an alternative philosophy which was to treat the in-control ARL as an explicit design parameter. The above method allows the rapid computation of the ARL for a range of values for the control limits. In Fig. 4, the ARL as a function of the control limit is shown for selected values of α. The control limit is given in terms of the standard deviation of y_t. From this figure the control limits for an AR(1) process can be determined.

Table 1 Control limits to guarantee an in-control ARL of 370 for AR(1) process

α	Control Limit	α	Control Limit
0	3.00	0.5	2.98
0.1	3.00	0.6	2.96
0.2	3.00	0.7	2.93
0.3	3.00	0.8	2.86
0.4	2.99	0.9	2.71

Since the calculations of the ARL closely match the results of the Monte Carlo simulation, Fig. 5, the impact of the mean shift on the ARL is investigated in greater detail without further comparisons being undertaken with simulations. In Fig. 5 the solid line refers to the theoretical calculations and the dots with arrow bars are the Monte Carlo simulations with the associated three standard errors of the mean

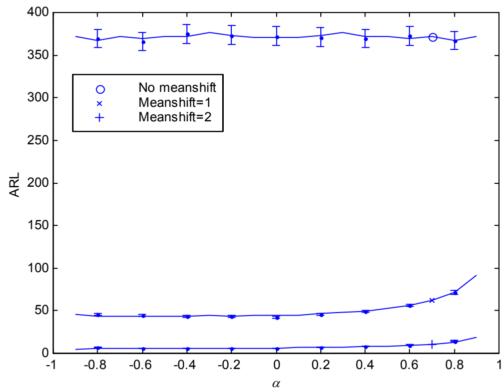

Fig. 5. The ARL as a function of alpha with adjusted control limits, with various mean shifts applied on the time series y_t.

The ARL of a Shewhart control chart with the control limits as given in Table 1 are shown in Fig. 5 where various mean shifts are applied to the time series data, y_t. In this situation, the control limit for negative α is equal to that of its positive counterpart. The calculations are in agreement with the Monte Carlo simulations. When no mean shift is applied, which corresponds to the in-control situation, the calculations do not give exactly 370, because of the rounding of the control limit to two decimals in Table 1. Compared with Fig 3, the in-control ARL does not deviate from 370.

In Fig. 6, the ARL is shown as a function of the mean shift and α. It can be observed that the ARL is only dependent on the mean shift and not on α, except for large positive values of α. Thus a Shewhart control chart with control limits adjusted to ensure the desired in-control ARL will exhibit the same sensitivity for equal sized mean shits regardless of the value of α. In practice the autoregressive parameter, α, is determined by either matching the autocorrelation function (Kaskavelis, 2000) or through estimation of the autoregressive parameter of an AR(1) process from the data.

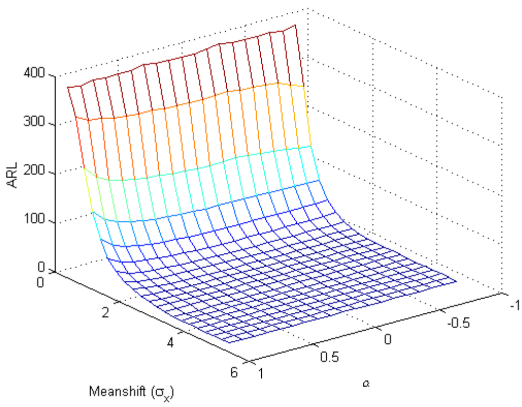

Fig. 6. The theoretically computed ARL as a function of alpha and the mean shift.

4. CONCLUSIONS

Within the paper, it is shown, based on the in-control probability at individual points in a control chart, how the density function of the run length of control charts can be determined. The density function can consequently be used to calculate the Average Run Length (ARL) of a control chart. The ARL is a widely used metric for comparing between monitoring strategies in SPC. The proposed approach is more generic than that described by Wetherill and Brown (1991).

The theoretical ARL for in-control data and out-of control data with step changes in the mean were calculated for three cases, independent data, the residuals of AR(1) models and serially correlated data. The theoretical results corresponded to the ARL obtained through Monte Carlo simulations. It is also shown that, in contrast to the claim of Schmid (1995), the ARL of serial correlated data in-control charts can be computed.

The in-control ARL's were computed as a function of the magnitude of the control limits. The control limits for Shewhart charts that realise an in-control ARL of 370 were determined for various values of autoregressive parameter for an AR(1) process. The impact of mean shifts on the performance of the ARL was subsequently investigated. It was found that the ARL depends only on the mean shift and not on α, except for large positive values of α.

The outcome of this work is that time consuming Monte Carlo simulations can now be replaced by the approach proposed for the assessment of the performance of control charts. This work can also be extended to more complicated SPC monitoring schemes, such as multivariate systems. However for multivariate problems, the problem is compounded by the fact that the parameter space may be large making the problem computationally intensive.

5. ACKNOWLEDMENTS

PM acknowledges the EU project PERFECT (Esprit Project 28870) for supporting his PhD studies.

6. REFERENCES

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

Apley, D.W. and Shi, J. (1999). The GLRT for statistical process control of autocorrelated processes. *IIE Transactions*, **31(12)**, pp. 1123-1134.

Box, G.E.P., G.M. Jenkins and G.C. Reinsel (1994). *Time series analysis: Forecasting and control*, Prentice Hall, Englewood Cliffs, NJ, USA.

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

Harris, T. J., and Ross, W. H. (1991). Statistical process-control procedures for correlated observations. *Canadian Journal of Chemical Engineering*, **69(1)**, pp. 48-57.

Kaskavelis, E. (2000). Statistical monitoring and prediction of petrochemical processes. *PhD, University of Newcastle, Newcastle upon Tyne, UK.*

Lowry, C. A., W.H. Woodall, C.W. Champ and S.E. Rigdon (1992). A multivariate exponentially weighted moving average control chart. *Technometrics*, **34(1)**, pp. 46-53.

Montgomery, D.C. (1991). *Introduction to statistical quality control*, John Wiley & Sons, Singapore.

Montgomery, D.C. and C.M. Mastrangelo (1991). Some statistical process-control methods for autocorrelated data. *Journal of Quality Technology*, **23(3)**, pp. 179-193.

Mulder, P., E.B. Martin and A.J. Morris (2002). The monitoring of dynamic processes using multivariate time series. *Advances in Process Control 6*, York, UK, pp. 194-201.

Papoulis, A. (1991). *Probability, random variables, and stochastic processes*, McGraw-Hill, New York.

Schmid, W. (1995). On the run-length of a shewhart chart for correlated data. *Statistical Papers*, **36(2)**, pp. 111-130.

Wardell, D. G., H. Moskowitz and R.D. Plante (1994). Run-length distributions of special-cause control charts for correlated processes. *Technometrics*, **36(1)**, pp. 3-17.

Wetherill, G.B. and D.W. Brown (1991). *Statistical process control: Theory and practice,* Chapman and Hall, London, UK.