Research Article

An Approximation to the Average Run Length on a CUSUM Control Chart with a Numerical Integral Equation Method for a Long-Memory ARFIMAX Model

Wilasinee Peerajit[1]*

[1]Department of Applied Statistics, Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, 1518 Pracharat 1 Road, Bangsue, Bangkok 10800, Thailand
*E-mail: wilasinee.p@sci.kmutnb.ac.th

Abstract

The CUSUM control chart performance based on the out-of-control average run length (ARL₁) can rapidly detect small-to-moderate-sized changes in the process mean. The main goal of this study is to approximate the ARL using a numerical integral equation (NIE) method for detecting changes in the mean on a cumulative sum (CUSUM) control chart for a long-memory process with exponential white noise. The target long-process observations are derived from an autoregressive fractionally integrated moving-average (ARFIMA) model with explanatory variable X (ARFIMAX) by focusing on X. The NIE method is derived by using the Gauss-Legendre quadrature rule technique, which is applied by solving a system of linear equations base on integral equations. In addition, approximation by this method was numerically evaluated and the results compared with an ARL based on explicit formulas in terms of the percentage relative error. It was found that the proposed NIE method had comparable performance with the explicit-formula-based ARL method. To demonstrate its practicality, the proposed NIE method was applied to the stock price data for the Siam Commercial Bank in Thailand with the explanatory variable as the exchange rate of Thai Baht against the US Dollar. The results reveal that it performed well in this practical implementation. Thus, the NIE method is a good alternative to using explicit formulas for constructing the ARL to detect changes in the mean for a long-memory process with exponential white noise on a CUSUM control chart.

Keywords: ARFIMAX(p,d,q,r), exponential white noise, numerical integral equation (NIE) method

Introduction

Control charts are tools frequently used in statistical process control, one of which is the cumulative sum (CUSUM) control chart that has received a great deal of attention for use in modern industrial processes. The CUSUM control chart has become a popular tool to detect small-to-moderate-sized changes in the process mean as well as to control the quality of products from manufacturing processes. Page (1954) initially conducted a study on the CUSUM control chart, followed by Gan (1991), Luceño & Puig-Pey (2006), Wu & Wang (2007).

A criterion for measuring the performance of a CUSUM control chart is the average run length (ARL). To efficiently monitor a process, ARL_0, which is the ARL when there is no shift in the process parameter (the in-control process), should be as large as possible, whereas ARL_1, which is the ARL when there is a shift in the process parameter (the out-of-control process), should be as small as possible. (cf. Riaz et al. (2014)). It has been found that the performance of a CUSUM control chart is significantly affected by autocorrelation (e.g., Johnson & Bagshaw, 1974; Lu & Reynolds, 2001; Kim et al., 2007). However, since this autocorrelation is often an inherent part of a process, it must be properly modeled and monitored.

Observations or actual data can be compiled from a random process based on a time series. In particular, some of these models have been derived based on econometric data. Both autoregressive (AR) and moving-average (MA) elements can be found in econometric observations following a time series. An important factor to be considered when establishing a model is how to measure the errors (the difference between the real values and the approximated ones). Smaller errors infer a more efficient model. In general, white noise (or normally distributed white noise) indicates the errors in a time series model with autocorrelated observations. However, the white noise can be distributed differently in some applications, such as exponentially distributed white noise. For instance, based on the study by Jacob & Lewis (1997), an ARMA (1,1) process order was indicated when the observations were considered as exponential white noise. Later, Mohamed & Hocine (2003) applied Bayesian principles to an AR(1) process with exponential white noise, while Pereira & Turkrman (2004) similarly applied exponential white noise in a Bayesian analysis of threshold AR models. Recently, the estimation of AR model parameters with exponential white noise for cases with an unknown order was introduced by Suparman (2018).

The traditional model used for short-memory processes such as $AR(p), MA(q), ARMA(p,q)$, and $ARIMA(p,d,q)$ cannot be applied to long-memory processes. This problem has been solved by the establishment of several models, a popular one being the AR fractionally integrated MA ($ARFIMA(p,d,q)$) model, as proposed by Granger & Joyeux (1980). In addition, ARFIMA with an explanatory variable X ($ARFIMAX(p,d,q,r)$) is an extension of ARFIMA. Granger & Joyeux (1980) and Hosking (1981) established models for long-memory processes with observed fluctuations. In addition, Ebens (1999) suggested ARFIMAX models for the determination of observed fluctuations in a Dow Jones Industrial Average portfolio and found correlation between the econometric models and economic indicators (variables) that have an impact on economic forecasting. Examples of explanatory variables, which are independent of other variables in the system, are government investment policies, exchange rates, interest rates, inflation rates, etc. These variables have an impact on the econometric model when forecasting the future economic situation. If there is an explanatory variable involved in any type of forecasting, the model's predictive power will be higher than forecasting without it.

There are many types of control charts associated with time series with and without ARFIMA modeling. For example, the correlated data presented in the performance of Shewhart and exponentially weighted MA (EWMA) control charts with ARFIMA model processes was investigated by Ramjee (2000). The study revealed that although process shifts were detected,

these charts were not efficient. Two years later, the forecast-based hyperbolic weighted MA (HWMA) chart was created by Ramjee et al. (2002) for non-stationary ARFIMA model processes. Pan and Chen (2008) monitored the air quality in Taiwan by using control charts for autocorrelated data based on ARFIMA and ARIMA model processes. Subsequently, they concluded that the control charts for ARFIMA processes are more efficient than for the ARIMA ones. Later, a EWMA control chart was proposed by Rabyk & Schmid (2016) to detect changes in the mean of a long-memory process. The control was derived from calculations applied for an ARFIMA(p,d,q) process. Meanwhile, the current study is focused on CUSUM control charts for detecting changes in the mean of an ARFIMAX model long-memory process.

Voluminous research in various areas has suggested that computation of the ARL should be used to measure the performance of control charts. The main methods for computing ARL in the literature are Monte Carlo simulation, the Markov chain approach, explicit formulas, and integral equations. The Markov chain approach was originally introduced by Brook & Evans (1972) to study the run length properties of a CUSUM control chart under the assumption of independent and identically distributed (i.i.d.) observations. Hawkins (1992) suggested improving the Markov chain method by using Richardson extrapolation for a whole distribution family, including the Chi-squared distribution. Champ & Rigdon (1991) proposed the integral equation approach for evaluating ARL. Fredholm integral equations of the second kind have been used as part of the NIE method applied to calculate the ARL (Wieringa, 1999). Recently, Acosta-Mejía et al. (1999) utilized the integral equation approach that was numerically approximated by using the Gauss-Legendre quadrature method. Note that the sample variance has a Chi-squared distribution that is skewed to the right and restricted on the half-real line. Knoth (2006) suggested using a piecewise collocation method to approximate ARL integral equations instead of using the Gauss-Legendre quadrature-based method that depends on numerical quadrature rules to measure the integrals. The numerical integration (or quadrature) method is commonly denoted as an approximation of the integral. Some basic quadrature rules used for the approximation of integration equations are the midpoint rule, the composite trapezoidal rule, composite Simpson's rule, and the Gaussian rule. Where midpoint rule is the alternative to approximation for ARL. Hence, approximating the ARL obtained from the integral equation technique and midpoint rule, which comprise the basis of the numerical integral equation (NIE) method (see Paichit, 2017; Peerajit et al., 2018) is of interest in this study.

In this study, an approximate ARL is constructed using the NIE method for use on a CUSUM control chart for a long-memory process with exponential white noise. The target long-process observations are derived based on the ARFIMAX model with a focus on the explanatory variable X. The rest of this paper is organized as follows. In the next section, background on the CUSUM control chart, the characteristics of the ARL, and the general ARFIMAX(p, d, q, r) process are introduced and proved. The approximated ARL for the CUSUM control chart by using an NIE method for a long-memory process is presented in Section 3. A comparison of the analytical results is given in Section 4. In Section 5, stock price data for Siam Commercial Bank (SCB) in Thailand with the explanatory variable as the exchange rate of Thai baht (THB) against the US dollar (USD) were used to evaluate the ARLs derived from applying the NIE method and the explicit formulas, while Section 6 offers conclusions on the study.

The CUSUM Control Chart for an ARFIMAX(p,d,q,r) Process with Exponential White Noise

1) The CUSUM Control Chart for Monitoring Shifts in the Process Mean

The concept of the CUSUM control chart was first proposed by Page (1954). It is well known that it performs better than the Shewhart control chart for small-to-medium-sized changes in the process mean. The commonly used form of the upper-sided CUSUM control chart is based on the sequence

$$C_t = \max\left\{C_{t-1} + Y_t - v,\, 0\right\}, \text{ for } t = 1, 2, \dots, \tag{1}$$

where C_t is the CUSUM control chart statistic used for detecting upward shifts, quantity C_{t-1} denotes the previous value of the statistic (its initial value C_0 is set to ψ and a parameter of process Y_t is defined as the generally formed sequence of an ARFIMAX(p,d,q,r) process with exponential white noise), and v is a suitably chosen positive constant termed the reference value.

Let τ_h be the stopping time for the CUSUM control chart with a predetermined threshold of h (a constant parameter called the decision limit value or upper control limit). Thus,

$$\tau_h = \inf\left\{t > 0;\, C_t > h\right\}, \text{ for } \psi < h, \tag{2}$$

Note that when $0 < C_t < h$, then the process is in-control whereas when $C_t > h$, then the process is out-of-control.

2) Characteristics of the ARL

Assume that sequential observations $\xi_1, \xi_2, \xi_3, \dots$ are a sequence of i.i.d. random variables with distribution function $F(\xi, \alpha)$, where α is a control parameter. Moreover, $\alpha = \alpha_0$ and $\alpha_1 > \alpha_0$ are before and after the change-point time $\upsilon \leq \infty$ respectively.

The ARL is defined as $E_\upsilon(.)$, which is the expectation that the change point from $\alpha = \alpha_0$ to $\alpha = \alpha_1$ for $F(\xi, \alpha)$ occurs at time υ, where $\upsilon \leq \infty$. Therefore, an appropriate control chart provides a large ARL for $\upsilon = \infty$ and small ARL for $\upsilon = 1$. There are two values for the ARL. The first is obtained by substituting α for the state of no change at all ($\alpha = \alpha_0$). This is the in-control state (ARL$_0$), which is the expectation of the change point τ_h defined as

$$ARL_0 = E_\infty(\tau_h) = \gamma, \tag{3}$$

where γ is a constant for ARL_0.

On the other hand, when α is substituted by the state of change already at the beginning ($\alpha_1 > \alpha_0$), the ARL is only evaluated for the special scenario of $\upsilon = 1$. Hence, the conditional for the out-of-control ARL process (ARL₁) is as follows:

$$\text{ARL}_1 = \text{E}_1(\tau_h \mid \tau_h \geq 1) \tag{4}$$

3) The General ARFIMAX Process on a CUSUM Control Chart

The ARFIMAX(p,d,q,r) model, where p is the order of the autoregressive (AR) process, d is the fractional integration parameter, q is the order of the MA process, and r is the explanatory variable order in the model (Ebens (1999)) can be written as

$$\phi_p(B)(1-B)^d(Y_t - \mu) = \sum_{i=1}^{r} \beta_i X_{it} + \theta_q(B)\xi_t, \tag{5}$$

where μ is the constant process mean of $\{Y_t\}$, X_{it} is an explanatory variable, β_i is an unknown parameters, ξ_t is a white noise process assumed to be exponentially distributed with $\xi_t \sim Exp(\alpha)$ when shift parameter $\alpha > 0$. The terms $\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - ... - \phi_p B^p$ and $\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - ... - \theta_q B^q$ are AR and MA polynomials on B (the backward-shift operator), i.e. $B^k Y_t = Y_{t-k}$ for the kth order. The fractional (d) can take on non-integer values in the range $(-0.5, 0.5)$ and represents the fractional order of integration.

For $-0.5 < d < 0.5$, the fractional white noise process is defined as

$$(1-B)^d Y_t = \xi_t.$$

The operator $(1-B)^d$ is defined as the fractional difference for $d \in (-0.5, 0.5)$. This can be determined naturally by using binomial series expansions obtained from

$$(1-B)^d := \sum_{k=0}^{\infty} \binom{d}{k}(-B)^p = 1 - dB + \frac{1}{2}d(d-1)B^2 - \frac{1}{6}d(d-1)(d-2)B^3 + ... , .$$

which was generalized by Granger & Joyeux (1980) and Hosking (1981). d in the fractionally integrated model can take on non-integer values but is often restricted to $|d| < 0.5$. In particular, the case of $0 < d < 0.5$ corresponds to a stationary and invertible model with long-memory representation, which is of interest in this study. This ARFIMAX model is typically analyzed as an ARMAX model where the innovations are fractionally integrated.

Note that, by rewriting Equation (5), the fractionally integrated process can be defined as

$$(1-B)^d Y_t = Y_t - dBY_t + \frac{1}{2}d(d-1)B^2 Y_t - \frac{1}{6}d(d-1)(d-2)B^3 Y_t + ... , = \xi_t$$

$$Y_t = dY_{t-1} - \frac{1}{2}d(d-1)Y_{t-2} + \frac{1}{6}d(d-1)(d-2)Y_{t-3} - ... + \xi_t ,$$

Therefore, the general ARFIMAX(p,d,q,r) process (Y_t) with exponential white noise on a CUSUM control chart is written as

$$\begin{aligned}
Y_t =\ & \mu + \beta_1 X_{1t} + \beta_2 X_{2t} + ... + \beta_r X_{rt} + \xi_t - \theta_1 \xi_{t-1} - \theta_2 \xi_{t-2} - ... - \theta_q \xi_{t-q} \\
& + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} + d(Y_{t-1} - \phi_1 Y_{t-2} - \phi_2 Y_{t-3} - ... - \phi_p Y_{t-p-1}) \\
& + \frac{1}{2}d(d-1)(-Y_{t-2} + \phi_1 Y_{t-3} + \phi_2 Y_{t-4} + ... + \phi_p Y_{t-p-2}) \\
& + \frac{1}{6}d(d-1)(d-2)(Y_{t-3} - \phi_1 Y_{t-4} - \phi_2 Y_{t-5} - ... - \phi_p Y_{t-p-3}) + ... ,
\end{aligned} \tag{6}$$

where ϕ_i is an autoregressive coefficient, where $|\phi_i| < 1$; $i = 1, 2, ..., p$, and θ_i is an MA coefficient; $|\theta_i| < 1$; $i = 1, 2, ..., q$. The initial value of an ARFIMAX(p, d, q, r) process requires that $Y_{t-1}, Y_{t-2}, ..., Y_{t-p}, Y_{t-(p+1)}, ..., \xi_{t-1}, \xi_{t-2}, ..., \xi_{t-q}$ and $X_{1t}, X_{2t}, ..., X_{rt}$ are equal to 1.

The NIE Method for Approximating the ARL for a CUSUM Control Chart
The goal of this section is to approximate the ARL with the NIE method using the Fredholm's integral equation of the second kind (Champ & Rigdon (1991)). From then on, the NIE was applied to the CUSUM control chart statistic to approximate the ARL. As is known, this Gauss-Legendre rule technique for numerical computation of an integral equation can be applied to the NIE method

1) The Integral Equation for a One-sided CUSUM Control Chart
Let P_c and E_c denote the probability measure and induced expectation corresponding to the initial value ψ, respectively. The function ARL(ψ) denotes the ARL of the ARFIMAX model when the CUSUM starts with initial value ψ, and the initial value of the monitoring CUSUM control chart statistic is $C_0 = \psi$; $0 < \psi < h$. Subsequently, $\text{ARL}(\psi) = \text{E}_\theta(\tau_h) < \infty$ satisfies the following integral equation:

$$\text{ARL}(\psi) = 1 + P_c\{C_1 = 0\}\,\text{ARL}(0) + E_c[I\{0 < C_1 < h\}\,\text{ARL}(C_1)], \tag{7}$$

where the first observation is C_1, for which $I\{0 < C_1 < h\}$ is the indicator function.

The integral equation for computing the ARL of a one-sided CUSUM control chart is derived from Equation (7) using a Fredholm integral equation of the second kind (Kharab & Guenther (2011)) such that ARL(ψ) can be rewritten as

$$\text{ARL}(\psi) = 1 + \text{ARL}(0)F\left(v - \psi - Y_t\right) + \int_0^h \text{ARL}(u)f(u + v - \psi - Y_t)du, \qquad (8)$$

where $f(.)$ and $F(.)$ are the probability density function (pdf) and cumulative distribution function (cdf) of ξ_t, respectively, and can be approximated numerically.

2) The Approximated ARL with the NIE Method Using the Gauss-Legendre Rule Technique

According to the integral equation in Equation (8), the Gauss-Legendre quadrature rule technique can be used to obtain numerical solutions. Obviously, integral $\int_0^h f(u)du$ can be approximated by the sum of areas of rectangles with bases h/m and heights chosen as the values of f at midpoints of intervals of length h/m beginning at zero. Interval $[0, h]$ provides the division points into $a_1 \le ... \le a_m$ and $w_j = h/m \ge 0$ is a set of weights. The approximation for the integral is obtained in the summation form as

$$\int_0^h W(u)f(u)du \approx \sum_{j=1}^m w_j f(a_j)$$

where $W(u)$ is a weight function, $a_j = \dfrac{h}{m}\left(j - \dfrac{1}{2}\right)$, and $w_j = h/m; j = 1, 2, ..., m$.

Let $\text{ARL}^{NIE}(\psi)$ denote the approximated ARL with NIE using the Gauss-Legendre rule technique on the interval $[0, h]$. The integral equation can be transformed as a system of algebraic linear equations: m equations in the m unknowns. Therefore, the integral equation in Equation (8) comprises the set $\text{ARL}^{NIE}(\psi) = \text{ARL}^{NIE}(a_1), ..., \text{ARL}^{NIE}(a_m)$, which can be approximately written as

$$\text{ARL}^{NIE}(a_1) = 1 + \text{ARL}^{NIE}(a_1)\left[F(v - a_1 - Y_t) + w_1 f(v - Y_t)\right] + \sum_{j=2}^m w_j \text{ARL}^{NIE}(a_j)f(a_j + v - a_1 - Y_t),$$

$$\text{ARL}^{NIE}(a_2) = 1 + \text{ARL}^{NIE}(a_1)\left[F(v - a_2 - Y_t) + w_1 f(a_1 + v - a_2 - Y_t)\right] + \sum_{j=2}^m w_j \text{ARL}^{NIE}(a_j)f(a_j + v - a_2 - Y_t),$$

$$\vdots$$

$$\text{ARL}^{NIE}(a_m) = 1 + \text{ARL}^{NIE}(a_1)\left[F(v - a_m - Y_t) + w_1 f(a_1 + v - a_m - Y_t)\right] + \sum_{j=2}^m w_j \text{ARL}^{NIE}(a_j)f(a_j + v - a_m - Y_t),$$

where Y_t is the general ARFIMAX(p, d, q, r) process for a CUSUM control chart.

This can be written in matrix form as

$$\mathbf{L}_{m \times 1} = \mathbf{1}_{m \times 1} + \mathbf{R}_{m \times m} \mathbf{L}_{m \times 1}, \tag{9}$$

where $\mathbf{L}_{m \times 1} = \left[\mathrm{ARL}^{NIE}(a_1), \mathrm{ARL}^{NIE}(a_2), ..., \mathrm{ARL}^{NIE}(a_m) \right]^T$, $\mathbf{1}_{m \times 1} = \left[1, 1, ..., 1 \right]^T$ is a column vector of $\mathrm{ARL}^{NIE}(a_j)$ and 1's, and $\mathbf{R}_{m \times m}$ is a matrix with dimension $m \times m$:

$$\mathbf{R}_{m \times m} = \begin{bmatrix} F(v - a_1 - Y_t) + w_1 f(v - Y_t) \ ... & +w_m f(a_m + v - a_1 - Y_t) \\ F(v - a_1 - Y_t) + w_1 f(a_1 + v - a_2 - Y_t) \ ... & +w_m f(a_m + v - a_2 - Y_t) \\ \vdots & \vdots \\ F(v - a_m - Y_t) + w_1 f(a_1 + v - a_m - Y_t) \ ... + w_m f(a_m + v - a_m - Y_t) \end{bmatrix}.$$

The matrix form of Equation (9) is equivalent to

$$(\mathbf{I}_m - \mathbf{R}_{m \times m}) \mathbf{L}_{m \times 1} = \mathbf{1}_{m \times 1}. \tag{10}$$

where $\mathbf{I}_m = diag(1, 1, ..., 1)$ is the unit matrix order m. If $(\mathbf{I}_m - \mathbf{R}_{m \times m})$ is invertible and exists, then the approximated solution with the NIE method for the integral equations in the matrix is as follows:

$$\mathbf{L}_{m \times 1} = (\mathbf{I}_m - \mathbf{R}_{m \times m})^{-1} \mathbf{1}_{m \times 1}, \tag{11}$$

where $\mathbf{L}_{m \times 1} = \left[\mathrm{ARL}^{NIE}(a_1), \mathrm{ARL}^{NIE}(a_2), ..., \mathrm{ARL}^{NIE}(a_m) \right]^T$ after computing $\mathrm{ARL}^{NIE}(a_1)$, $\mathrm{ARL}^{NIE}(a_2)$, ..., $\mathrm{ARL}^{NIE}(a_m)$, from which $\mathrm{ARL}^{NIE}(\psi)$ can be obtained. Subsequently, the NIE for $\mathrm{ARL}^{NIE}(\psi)$ can be written as

$$\mathrm{ARL}^{NIE}(\psi) = 1 + \mathrm{ARL}^{NIE}(a_1) F(v - \psi - Y_t) + \sum_{j=1}^{m} w_j \mathrm{ARL}^{NIE}(a_j) f(a_j + v - \psi - Y_t), \tag{12}$$

Thus,

$$\mathrm{ARL}^{NIE}(\psi) = 1 + \mathrm{ARL}^{NIE}(a_1) F(v - \psi - \mu - \beta_1 X_{1t} - ... - \beta_r X_{rt} - \xi_t + \theta_1 \xi_{t-1} + ... + \theta_q \xi_{t-qt}$$

$$-\phi_1 Y_{t-1} - ... - \phi_p Y_{t-p} - d(Y_{t-1} + \phi_1 Y_{t-2} + .. + \phi_p Y_{t-p-1}) - \frac{1}{2} d(d-1)(-Y_{t-2} + \phi_1 Y_{t-3} + ... + \phi_p Y_{t-p-2})$$

$$-\frac{1}{6} d(d-1)(d-2)(Y_{t-3} - \phi_1 Y_{t-4} - .. - \phi_p Y_{t-p-3}) + ...)$$

$$+\sum_{j=1}^{m} w_j \mathrm{ARL}^{NIE}(a_j) f(a_j + v - \psi - \mu - \beta_1 X_{1t} - ... - \beta_r X_{rt} - \xi_t + \theta_1 \xi_{t-1} + ... + \theta_q \xi_{t-qt})$$

$$-\phi_1 Y_{t-1} - ... - \phi_p Y_{t-p} - d(Y_{t-1} + \phi_1 Y_{t-2} + .. + \phi_p Y_{t-p-1}) - \frac{1}{2} d(d-1)(-Y_{t-2} + \phi_1 Y_{t-3} + ... + \phi_p Y_{t-p-2})$$

$$-\frac{1}{6} d(d-1)(d-2)(Y_{t-3} - \phi_1 Y_{t-4} - .. - \phi_p Y_{t-p-3}) + ...), \tag{13}$$

with $w_j = \dfrac{h}{m}$, and $a_j = \dfrac{h}{m}\left(j - \dfrac{1}{2}\right)$; $j = 1, 2, \ldots, m.$

This is the new approximated ARL using the NIE method on a CUSUM control chart for a long-memory process with exponential white noise. Hence, the conventional Gauss-Legendre quadrature rule technique can be used to approximate the ARL quite accurately.

Numerical Results

The performance of the approximated ARL using the NIE method on a CUSUM control chart for detecting changes in the mean of a long-memory ARFIMAX process was compared with the analytical ARL. Explanatory variable (X) and long-memory processes $d = 0.1, 0.2$ and 0.3 were considered. In addition, the Gauss-Legendre quadrature rule module in the Mathematica program with the number of division points $m = 700$ nodes was used to calculate the numerical results. Furthermore, the analytical ARL for the explicit formulas can be written as

$$\mathrm{ARL}^{EF}(\psi) = e^{\alpha h}(1 + e^{\alpha(v - Y_t)} - \alpha h) - e^{\alpha \psi}. \tag{14}$$

where Y_t is the general $\mathrm{ARFIMAX}(p, d, q, r)$ process for a CUSUM control chart.

The criterion for consideration to evaluate the performance of the methods is the percent relative error $(\varepsilon(\%))$, which is the magnitude of the difference between the expressions for the ARL using the NIE method ($\mathrm{ARL}^{NIE}(\psi)$) (see Equation 13) and the analytical ARL of the explicit formulas ($\mathrm{ARL}^{EF}(\psi)$) (see Equation 14) divided by the magnitude of ARL^{EF} as a percentage:

$$\varepsilon(\%) = \left| \frac{\mathrm{ARL}^{EF}(\psi) - \mathrm{ARL}^{NIE}(\psi)}{\mathrm{ARL}^{EF}(\psi)} \right| \times 100\%, \tag{15}$$

when $\varepsilon(\%)$ is less than 0.30%, the results obtained from the two methods are similar and in excellent agreement.

The parameter settings for the CUSUM control chart that result in the desired in-control ARL and minimize the out-of-control ARL (ARL$_1$) are needed for the specified changes in the process. The process follows an exponential distribution with mean parameter (α) when it is in-control ($\alpha_0 = 1$). On the other hand, the out-of-control process for parameter $\alpha_1 > 1$ is in the format $\alpha_1 = (1 + \delta)\alpha_0$, where δ is the magnitude of change in the process mean ($\delta = 0.01, 0.03, 0.05, 0.10, 0.30, 0.50, 0.80, 1$, and 3). When considering the parameter values of v and h for an upper-sided CUSUM control chart, the reference values $v = 3.0, 3.5, 4.0$, and 4.5 for the CUSUM control limit are used and the decision value (h) is chosen to achieve the specified ARL$_0$ = 370.

Table 1. Values of CUSUM control limit (h) with corresponding values of ν for the desired $ARL_0 = 370$ for ARFIMAX(p,d,q,r) models.

Parameters of models				Long-memory	ν			
ϕ_1	ϕ_2	θ_1	β_1	with models	3.0	3.5	4.0	4.5
0.1	-	0.1	0.75	[1] ARFIMAX(1, 0.1, 1, 1)	4.361765	3.537212	2.917920	2.365228
				[2] ARFIMAX(1, 0.2, 1, 1)	4.711590	3.733592	3.080743	2.515072
				[3] ARFIMAX(1, 0.3, 1, 1)	5.209625	3.922684	3.229655	2.649611
0.1	0.2	0.1	0.75	[4] ARFIMAX(2, 0.1, 1, 1)	4.794150	3.772573	3.112120	2.543635
				[5] ARFIMAX(2, 0.2, 1, 1)	5.595760	3.994427	3.283828	2.697874
			0.50	[6] ARFIMAX(2, 0.3, 1, 1)	4.653225	3.704270	3.056923	2.493320

Table 1 shows the results for the computed h value for ARL_0 = 370. It was found that as ν increases, h decrease on the CUSUM control chart for long-memory process models [1]–[6]; e.g. for ν = 3, h = 4.361765, 4.711590, 5.209625, 4.794150, 5.59576, and 4.653225, respectively.

Table 2. Comparison of ARL_1 values using the NIE method and explicit formulas for an ARFIMAX(1, d, 1, 1) process on a CUSUM control chart for $ARL_0 = 370$.

δ	ν	3.0			3.5			4.0			4.5		
	h	4.361765	4.71159	5.209625	3.537212	3.733592	3.922684	2.91792	3.080743	3.229655	2.365228	2.515072	2.649611
	Model	[1]	[2]	[3]	[1]	[2]	[3]	[1]	[2]	[3]	[1]	[2]	[3]
0.01	ARL^{NIE}	344.140	342.763	**339.992**	346.640	345.896	**345.445**	347.284	347.667	**346.845**	347.853	347.721	**347.589**
	ARL^{Exp}	345.013	343.620	340.740	347.101	346.720	346.291	347.968	347.784	347.589	348.419	348.320	348.218
	$\varepsilon(\%)$	0.25	0.25	0.22	0.13	0.24	0.24	0.20	0.03	0.21	0.16	0.17	0.18
0.03	ARL^{NIE}	300.497	296.983	**289.850**	305.910	304.902	**303.782**	308.252	307.744	**307.216**	309.533	309.245	**308.951**
	ARL^{Exp}	301.230	297.689	290.430	306.594	305.607	304.503	308.845	308.364	307.859	310.026	309.766	309.498
	$\varepsilon(\%)$	0.24	0.24	0.20	0.22	0.23	0.24	0.19	0.20	0.21	0.16	0.17	0.18
0.05	ARL^{NIE}	263.777	258.784	**248.699**	271.493	270.054	**268.656**	274.831	274.108	**273.356**	276.641	276.236	**275.822**
	ARL^{Exp}	264.396	259.369	249.150	272.084	270.662	269.308	275.347	274.647	273.914	277.072	276.692	276.299
	$\varepsilon(\%)$	0.23	0.23	0.18	0.22	0.22	0.24	0.19	0.20	0.20	0.16	0.16	0.17
0.10	ARL^{NIE}	194.539	187.711	**174.151**	205.291	203.266	**201.028**	210.025	208.996	**207.928**	212.607	212.030	**211.439**
	ARL^{Exp}	194.955	188.086	174.394	205.710	203.693	201.458	210.398	209.384	208.327	212.921	212.391	211.785
	$\varepsilon(\%)$	0.21	0.20	0.14	0.20	0.21	0.21	0.18	0.19	0.19	0.15	0.17	0.16
0.30	ARL^{NIE}	73.683	68.031	**57.523**	83.303	81.406	**79.365**	87.917	86.889	**85.837**	90.574	89.969	**89.357**
	ARL^{Exp}	73.793	68.113	57.535	83.436	81.541	79.493	88.044	87.018	85.969	90.684	90.085	89.476
	$\varepsilon(\%)$	0.15	0.12	0.02	0.16	0.17	0.16	0.14	0.15	0.15	0.12	0.13	0.13
0.50	ARL^{NIE}	37.043	33.646	**27.672**	43.223	41.960	**40.624**	46.413	45.686	**44.953**	48.340	47.895	**47.448**
	ARL^{Exp}	37.082	33.670	27.661	43.278	42.014	40.675	46.468	45.742	45.001	48.390	47.946	47.500
	$\varepsilon(\%)$	0.11	0.07	0.04	0.13	0.13	0.13	0.12	0.12	0.11	0.10	0.11	0.11
0.80	ARL^{NIE}	18.295	16.676	**14.019**	21.498	20.814	**20.114**	23.327	22.875	**22.473**	24.469	24.198	**23.931**
	ARL^{Exp}	18.306	16.684	14.027	21.519	20.833	20.124	23.331	22.908	22.487	24.490	24.217	23.946
	$\varepsilon(\%)$	0.06	0.05	0.06	0.10	0.09	0.05	0.02	0.14	0.06	0.09	0.08	0.06
1.00	ARL^{NIE}	13.113	12.078	**10.457**	15.279	14.804	**14.319**	16.568	16.263	**15.961**	17.422	17.219	**17.018**
	ARL^{Exp}	13.120	12.080	10.451	15.291	14.815	14.330	16.581	16.275	15.974	17.434	17.231	17.031
	$\varepsilon(\%)$	0.05	0.02	0.06	0.08	0.08	0.08	0.08	0.08	0.08	0.07	0.07	0.08
3.00	ARL^{NIE}	3.469	3.432	**3.418**	3.627	3.583	**3.543**	3.770	3.734	**3.699**	3.890	3.860	**3.832**
	ARL^{Exp}	3.469	3.432	3.418	3.627	3.583	3.544	3.772	3.734	3.699	3.890	3.860	3.832
	$\varepsilon(\%)$	0.00	0.00	0.00	0.00	0.00	0.03	0.05	0.00	0.00	0.00	0.00	0.00

The bold figures are the minimum ARL_1 according to the shifts.
[1] ARFIMAX(1, 0.1, 1, 1), [2] ARFIMAX(1, 0.2, 1, 1) and [3] ARFIMAX(1, 0.3, 1, 1).

Table 3. Comparison of ARL_1 values using the NIE method and explicit formulas for an ARFIMAX$(2, d, 1, 1)$ process on a CUSUM control chart for $ARL_0 = 370$.

	v	3.0			3.5			4.0			4.5		
δ	h	4.79415	5.59576	4.653225	3.772573	3.994427	3.70427	3.11212	3.283828	3.056923	2.543635	2.6978735	2.49332
	Model	[4]	[5]	[6]	[4]	[5]	[6]	[4]	[5]	[6]	[4]	[5]	[6]
0.01	ARL^{NIE}	342.380	**336.764**	343.019	345.808	**345.258**	345.961	347.022	**346.759**	347.105	347.694	**347.539**	347.741
	ARL^{Exp}	343.228	337.315	343.882	346.637	346.111	346.781	347.745	347.513	347.812	348.300	348.178	348.336
	$\varepsilon(\%)$	0.25	0.16	0.25	0.24	0.25	0.24	0.21	0.22	0.20	0.17	0.18	0.17
0.03	ARL^{NIE}	296.002	**281.504**	297.639	304.683	**303.315**	305.062	307.637	**307.008**	307.822	309.185	**308.837**	309.29
	ARL^{Exp}	296.695	281.872	298.352	305.393	304.04	305.765	308.262	307.659	308.438	309.712	309.392	309.807
	$\varepsilon(\%)$	0.23	0.13	0.24	0.23	0.24	0.23	0.20	0.21	0.20	0.17	0.18	0.17
0.05	ARL^{NIE}	257.393	**236.968**	259.714	269.742	**267.790**	270.283	273.957	**273.060**	274.22	276.152	**275.661**	276.299
	ARL^{Exp}	257.964	237.204	260.307	270.353	268.411	270.888	274.5	273.624	274.756	276.613	276.145	276.751
	$\varepsilon(\%)$	0.22	0.10	0.23	0.23	0.23	0.22	0.20	0.21	0.20	0.17	0.18	0.16
0.10	ARL^{NIE}	185.823	**158.660**	188.976	202.828	**200.098**	203.587	208.781	**207.507**	209.155	211.910	**211.208**	212.119
	ARL^{Exp}	186.182	158.705	189.360	203.256	200.528	204.013	209.171	207.911	209.541	212.245	211.559	212.448
	$\varepsilon(\%)$	0.19	0.03	0.20	0.21	0.21	0.21	0.19	0.19	0.18	0.16	0.17	0.15
0.30	ARL^{NIE}	66.517	**46.289**	69.056	81.001	**78.529**	81.706	86.676	**85.427**	87.047	89.845	**89.120**	90.062
	ARL^{Exp}	66.589	46.21	69.144	81.137	78.656	81.838	86.806	85.560	87.176	89.961	89.241	90.177
	$\varepsilon(\%)$	0.11	0.17	0.13	0.17	0.16	0.16	0.15	0.16	0.15	0.13	0.14	0.13
0.50	ARL^{NIE}	32.760	**21.625**	34.251	41.695	**40.086**	42.157	45.538	**44.670**	45.798	47.804	**47.276**	47.963
	ARL^{Exp}	32.779	21.573	34.278	41.748	40.135	42.211	45.593	44.726	45.853	47.855	47.329	48.014
	$\varepsilon(\%)$	0.06	0.24	0.08	0.13	0.12	0.13	0.12	0.13	0.12	0.11	0.11	0.11
0.80	ARL^{NIE}	16.272	**11.539**	16.958	20.671	**19.826**	20.919	22.801	**22.305**	22.951	24.141	**23.822**	24.238
	ARL^{Exp}	16.276	11.515	16.967	20.690	19.843	20.938	22.822	22.327	22.972	24.161	23.843	24.258
	$\varepsilon(\%)$	0.02	0.21	0.05	0.09	0.09	0.09	0.09	0.10	0.09	0.08	0.09	0.08
1.00	ARL^{NIE}	11.823	**8.991**	12.255	14.706	**14.130**	14.876	16.201	**15.847**	16.309	17.177	**16.942**	17.249
	ARL^{Exp}	11.824	8.977	12.259	14.717	14.140	14.888	16.214	15.859	16.322	17.190	16.955	17.262
	$\varepsilon(\%)$	0.01	0.16	0.03	0.07	0.07	0.08	0.08	0.08	0.08	0.08	0.08	0.08
3.00	ARL^{NIE}	3.435	**3.427**	3.437	3.574	**3.529**	3.589	3.729	**3.685**	3.739	3.853	**3.820**	3.864
	ARL^{Exp}	3.435	3.427	3.437	3.575	3.529	3.590	3.727	3.686	3.740	3.854	3.821	3.865
	$\varepsilon(\%)$	0.00	0.00	0.00	0.03	0.00	0.03	0.05	0.03	0.03	0.03	0.03	0.03

The bold figures are the minimum ARL_1 according to the shifts.
[4] ARFIMAX(2, 0.1, 1, 1), [5] ARFIMAX(2, 0.2, 1, 1) and [6] ARFIMAX(2, 0.3, 1, 1).

Tables 2 and 3 summarize the comparison of ARL results using the two methods for various magnitudes of change in the process mean. The first column in the two tables lists the magnitude of change (δ) divided into two levels: small changes $(\delta = 0.01 - 0.30)$ and moderate changes $(\delta = 0.50 - 3.00)$, which both showed decreases in ARL. The out-of-control ARL using both methods decreased. In addition, both methods could detect small changes rapidly. The results show that the ARL value for the NIE method approached that of the explicit formulas and because of that, both methods obtained percent relative errors $(\varepsilon(\%))$ of less than 0.30%, showing that they are in excellent agreement when applying this metric. It is obvious that the NIE method and explicit formulas performed slightly differently for the fractional integration parameter $(d = 0.1, 0.2, 0.3)$ based on ARFIMAX$(1, d, 1, 1)$ and ARFIMAX$(2, d, 1, 1)$ on the CUSUM control chart. Furthermore, $\varepsilon(\%)$ decreased when the shift size increased.

In the study of this process, we considered explanatory variable (X) and long-memory processes $d = 0.1, 0.2, 0.3$.

From the results in Tables 2 and 3, the notable outcomes are as follows:

1. The ARL values of the NIE method ware similar to explicit formulas for detecting changes in the process mean, with decreasing order being [3] ARFIMAX(1, 0.3, 1, 1), [2] ARFIMAX(1, 0.2, 1, 1) and [1] ARFIMAX(1, 0.1, 1, 1). When analyzing the performance of a CUSUM control chart, the ARL value needs to be as small as possible. In this study, the performances with models [1] and [2] were slightly lower than for model [3] at each reference value level (see Tables 2).

2. In Table 3, the ARL values obtained from model [5] ARFIMAX(2, 0.2, 1, 1) indicate a better performance than the model [4] ARFIMAX(2, 0.1, 1, 1) when $\beta_1 = 0.75$ and model [6] ARFIMAX(2, 0.2, 1, 1) when $\beta_1 = 0.50$ at each reference value level.

3. As the value for parameter v increased, the ARLs obtained from the two methods decreased.

In summary, approximating the ARL by using the NIE method is a good alternative to explicit formulas for measuring the performance of detecting a change in the process mean on a CUSUM control chart because the results show that its performance was similar to and in excellent agreement with the explicit formula method. Thus, equivalent numerical accuracy was achieved using the NIE method.

Implementing the Upper-Sided CUSUM Control Chart for the Mean: An Example Using Real Data

In this section, the ARL is approximated by using the NIE method and explicit formulas to illustrate the practical application of our proposed CUSUM control chart. The real data are the stock prices for SCB in Thailand with the explanatory variable as the exchange rate of THB against USD (unit: Baht rate). The observations were collected daily (5 days per week) from February 2020 to June 2020, consisting of 100 observations (source: https://th.investing.com). The dataset was analyzed and fitted to the long-memory ARFIMAX$(1, 0.499998, 1, 1)$ process with coefficients $\phi_1 = 0.907764$, $\theta_1 = -0.468244$, and $\beta_1 = -5.883148$, as well as white noise significantly exponentially distributed with mean $(\alpha_0) = 2.0136$. For the one-sided control limit for the CUSUM control chart, reference value $v = 3$ and upper control limit values $(h) = 4.53619$ and 5.15611 were used for $\text{ARL}_0 = 370$ and 500, respectively. The ARLs on the CUSUM control chart were derived using the two methods, the results of which are summarized in Table 4. The results are obviously in agreement with those in Tables 2 and 3. The numerical results obtained from the NIE method approached the explicit formulas for all cases when detecting small-to-moderate-sized changes in the process mean. To sum up, the NIE method is a good alternative for approximating the ARL for real applications on a CUSUM control chart.

Table 4. Comparison of ARL_1 values using the NIE method and explicit formulas for a long-memory ARFIMAX$(1, 0.499998, 1, 1)$ process under data on stock prices for SCB with the explanatory (exchange rate of THB against USD) for in-control process $\alpha_0 = 2.0136$.

δ	$ARL_0 = 370$			$ARL_0 = 500$		
	$v = 3$, $h = 4.53619$		$\varepsilon(\%)$	$v = 3$, $h = 5.15611$		$\varepsilon(\%)$
	ARL^{NIE}	ARL^{Exp}		ARL^{NIE}	ARL^{Exp}	
0.01	358.638	359.128	0.14	483.780	484.530	0.15
0.03	338.168	338.626	0.14	454.726	455.426	0.15
0.05	319.226	319.653	0.13	427.925	428.577	0.15
0.10	277.690	278.052	0.13	369.461	370.009	0.15
0.30	168.771	168.970	0.12	218.583	218.876	0.13
0.50	110.955	111.074	0.11	140.469	140.640	0.12
0.80	66.100	66.162	0.09	81.359	81.445	0.11
1.00	49.553	49.596	0.09	60.044	60.103	0.10
3.00	9.918	9.922	0.04	11.038	11.043	0.05

Conclusions

An NIE method for approximating the ARL of a shift in the mean of a long-memory ARFIMAX model on a CUSUM control chart is proposed in this article. The method approximated the ARL quite accurately and is a good alternative for verifying explicit formulas for constructing the ARL to detect changes in the mean for a long-memory process, and thus we suggest that it can be successfully applied to real applications such as in economics, finance, agriculture, hydrology, environmental, etc. Its applicability to real observations that follow a long-memory ARFIMAX process with a seasonal variation (SARFIMAX) as well as exponential white noise is a subject for future studies. In future studies, the NIE method for approximating the ARL could be developed for observations that correspond to the exponential family with continuous and increasing functions, such as gamma and Weibull distributions, among others.

Acknowledgement

References

Acosta-Mejía, C.A., Pignatiello, Jr., J.J. & Rao, B.V. (1999). A comparison of control charting procedures for monitoring process dispersion. *IIE Transactions*, 31, 569–579.

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

Champ, C.W. & 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 Simulation and Computation*, 20, 191-204.

Ebens H. (1999). Realized stock index volatility. Working Paper No. 420, Department of Economics, Johns Hopkins University.

Gan F.F. (1991). An Optimal Design of CUSUM Quality Control Charts. *Journal of Quality Technology*, 23, 279-286.

Granger C.W.J. & Joyeux R. (1980). An introduction to long memory time series models and fractional differencing. *Journal of Time Series Analysis*, 1(1): 15-29.

Hawkins, D.M. (1992). Evaluation of average run lengths of cumulative sum charts for an arbitrary data distribution. *Communications in Statistics Simulation and Computation*, 21, 1001–1020.

Hosking, J.R.M. (1981). Fractional differencing. *Biometrika*, 68(1), 165-176.

Jacob P.A. & Lewis PAW. (1997). A mixed autoregressive-moving average exponential sequence and point process (EARMA 1,1). *Advances in Applied Probability*, 9(1), 87-104.

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

Kharab, A. & Guenther, R. B. (2011). *An Introduction to Numerical Methods: A MATLAB Approach* (3rd ed.). New York, USA: Chapman and Hall/CRC.

Kim, S.-H., Alexopoulos, C., Tsui, K.-L., & Wilson, J. R. (2007). A distribution-free tabular CUSUM chart for autocorrelated data. *IIE Transactions*, 39, 317-330.

Knoth, S. (2006). Computation of the ARL for CUSUM-S2 schemes. *Computational Statistics & Data Analysis*, 51, 499-512.

Lu, C.-W. & Reynolds, M.R. (2001). Cusum charts for monitoring an autocorrelated process. *Journal of Quality Technology*, 33(3), 316–334.

Luceño, A. & Puig-Pey, J. (2002). An accurate algorithm to compute the run length probability distribution, and its convolutions, for a CUSUM chart to control normal mean. *Computer Statistics and Data Analysis*, 38(3), 249-261.

Luceño A. & Puig-Pey J. (2006). The Random Intrinsic Fast Initial Response of One-sided CUSUM Charts. *Journal of Applied Statistics*. 33: 189-201.

Mohamed, I. & Hocine, F. (2003). Bayesian Estimation of an AR(1) Process with Exponential White Noise. *A Journal of Theoretical and Applied Statistics*. 37, 365-372.

Page, E.S. (1954). Continuous Inspection Schemes. *Biometrika*. 41, 100-114.

Pan, J.N. & Chen, S.T. (2008). Monitoring Long-memory Air Quality Data Using ARFIMA Model. *Environmetrics*, 19, 209-219.

Paichit, P. (2017). Exact expression for average run length of control chart of ARX(p) procedure. *KKU Science Journal*, 45(4), 948-958.

Peerajit W., Areepong Y. & Sukparungsee S. (2018). Numerical integral equation method for ARL of CUSUM chart for long-memory process with non-seasonal and seasonal ARFIMA models. *Thailand Statistician*, 16(1), 26-37.

Pereira, I.M.S. & Turkrman, M.A. (2004). Bayesian Prediction in Threshold Autoregressive Models with Exponential White Noise. *Sociedad de Estadistica E Investigacion Operativa Test,* 13, 45-64.

Rabyk, L. & Schmid, W. (2016). EWMA Control Charts for Detecting Changes in the Mean of a Long-memory Process. *Metrika*. 79, 267–301.

Ramjee, R. (2018). *Quality control charts and persistent processes.* (Ph.D. Thesis) Stevens Institute of Technology, Hoboken, New Jersey.

Ramjee, R., Crato, N. & Ray, B.K. A. (2002). Note on Moving Average Forecasts of Long Memory Processes with an Application to Quality Control. *International Journal of Forecasting*, 18, 291-297.

Riaz M. Abbasi S.A., Ahmad S. & Zaman B. (2014). On efficient phase II process monitoring charts. *The International Journal of Advanced Manufacturing Technology*, 70(9-12), 2263–2274

Suparman S. (2018). A new estimation procedure using a reversible jump MCMC algorithm for AR models of exponential white noise. *International Journal of GEOMATE*, 15(49), 85-91.

Wieringa, J.E. (1999). *Statistical process control for serially correlated data* (Ph.D. Thesis). University of Groningen, Netherlands.

Wu Z. & Wang Q.N. (2007). A Single CUSUM Chart Using a Single Observation to Monitor a Variable. *International Journal of Production Research*, 45, 719-741.