Songklanakarin J. Sci. Technol.
39 (6), 779-786, Nov. - Dec. 2017

http://www.sjst.psu.ac.th

Original Article

Analytical explicit formulas of average run length for long memory process with ARFIMA model on CUSUM control chart

Wilasinee Peerajit, Yupaporn Areepong*, and Saowanit Sukparungsee

Department of Applied Statistics, Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, Bang Sue, Bangkok, 10800 Thailand

Received: 26 May 2016; Revised: 14 August 2016; Accepted: 18 September 2016

Abstract

This paper proposes the explicit formulas for the derivation of exact formulas from Average Run Lengths (ARLs) using integral equation on CUSUM control chart when observations are long memory processes with exponential white noise. The authors compared efficiency in terms of the percentage of absolute difference to a similar method to verify the accuracy of the ARLs between the values obtained by the explicit formulas and numerical integral equation (NIE) method. The explicit formulas were based on Banach fixed point theorem which was used to guarantee the existence and uniqueness of the solution for ARFIMA(p,d,q). Results showed that the two methods are similar in good agreement with the percentage of absolute difference at less than 0.23%. Therefore, the explicit formulas are an efficient alternative for implementation in real applications because the computational CPU time for ARLs from the explicit formulas are 1 second preferable over the NIE method.

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

1. Introduction

CUSUM control chart was first introduced by Page (1954) and has been continually developed by many researchers e.g. (Bissell, 1969; Ewan, 1963; Ewan & Kemp, 1960; Hawkins & Olwell, 1998; Johnson & Leone, 1962; Lucas, 1976; Ryan, 1989). These are commonly used instead of the Shewhart chart as they directly incorporate all of the information in the sequence from the values and detect small shifts in the mean more quickly and can widely implement control processes. Usually, this involves an evaluation of the control chart performance based on the Average Run Lengths (ARLs).

Average Run Length (ARL) is the expected number of observations taken from an in-control process until the control chart falsely signals out-of-control. ARL, as a common characteristic, is widely used as a measure of performance of a control chart. Ideally, the ARL is large enough to keep the level of false alarms at an acceptable level. ARL_0 is the notation for the in-control Average Run Length. The out-of-control Average Run Length is denoted by Average Delay for the out-of-control process (ARL_1). It is defined as the expectation of delay time for a true alarm. This time should minimize the quantity as possible.

The ARLs have been widely applied to techniques of control charts, percentage points have also been recommended, for example (Barnard, 1954; Bissell, 1969). Evaluations of ARLs for the CUSUM control charts, for example (Brook & Evans, 1972; Ewan & Kemp, 1960; Fellner, 1990; Gan, 1992; Goel & Wu, 1971; Hawkins, 1992; Luceño & Puig-Pey, 2000, 2002; Page, 1954; Woodall, 1983) have been conducted.

The integral equation is encountered in a variety of applications from many fields including continuum mechanics, mathematical economics, queuing theory, potential theory, geophysics, electricity and magnetism,

* Corresponding author.
Email address: yupaporn.a@sci.kmutnb.ac.th

optimization, optimal control systems, communication theory, population genetics, medicine etc. The integral equation was provided by Page (1954) and was used to approximate the ARLs of control chart by assuming a small shift in mean. A computation program based on the integral equation procedure was given by Vance (1986). Goel and Wu (1971) provided a nomogram for the determination of chart parameters of a CUSUM control chart. Lashkari and Rahim (1982) and Chung (1992) reported economic designs of CUSUM control charts.

The model of autoregressive fractionally integrated moving average (ARFIMA) processes have a fractional differencing parameter (d) which are used to model a long memory (or so-called long range) and stationary and invertible when values of d take between (-0.5, 0.5). These processes were introduced by Granger and Joyeux (1980) and Hosking (1981); a detailed description of long memory processes can be found in e.g. (Baillie, 1996; Baran, 1994; Palma, 2007; Proietti, 2014). The long memory process is involved in a number of applications including finance and economics, environmental, science and engineering. Control charts have been used to combine the long memory process with time series. The control chart is necessary as it is a number of time series following the ARFIMA model. Papers by Ramjee (2000) also analyzed the performance of Shewhart and EWMA control charts for the presence of correlated data which occurred from an ARFIMA model. The study result showed that these charts cannot perform well when detecting process shifts. Hence, a new type of control chart and Hyperbolic Weighted Moving Average (HWMA) control chart was proposed. Two years later, Ramjee *et al.* (2002) introduced a HWMA forecast-based control chart, specially designed for non-stationary ARFIMA models. Caballero *et al.* (2002) performed a number of tests on the analysis of daily time series of mid-latitude near-surface air temperature by plotting long-range dependent processes. Furthermore, Pan and Chen (2008) studied control charts for autocorrelated data using ARFIMA model to monitor the long memory air quality data for comparison. The result showed that residual control charts using ARFIMA models were more appropriate than using ARIMA models. Recently, Rabyk and Schmid (2016) introduced EWMA control charts to detect changes in the mean of a long-memory process.

Exponential white noise coordinated with time series has also been investigated. Jacob and Lewis (1977) analyzed autoregressive moving average process order (1,1) denoted by ARMA(1,1) when observations are exponentially distributed with exponential white noise. The exponential white noise was also used Bayesian methods to analyze the autoregressive model as proposed by Mohamed and Hocine (2010).

Several techniques to evaluate ARLs for the CUSUM and EWMA control charts including Monte Carlo simulations (MC), Markov Chain approach (MCA), numerical integral equation (NIE) method and explicit formulas have been proposed in the previous literature. For example, The Markov Chain approach (MCA) was introduced by Brook and Evans (1972), and many researchers have studied this matter. In particular, Champ and Woodall (1987), and Champ and Rigdon (1991), Gan (1992) and Gan (1996) presented an accurate NIE method based on an integral equation to compute the ARLs of CUSUM charts under linear trends. Recently, Areepong (2009) proposed analytical derivation to find explicit formulas of ARLs for EWMA control charts when observations are exponentially distributed. For example, problems from mathematical explicit formulas of ARLs using Fredholm integral equation for one-sided EWMA control chart with Laplace distribution and CUSUM control chart with hyper-exponential distribution were presented by Mititelu *et al.* (2010).

Busaba *et al.* (2012) analyzed the explicit formulas of ARLs for CUSUM control chart in cases of stationary first order autoregressive; AR(1), process with exponential white noise. The numerical integral equation (NIE) method of ARLs using the Gauss-Legendre numerical integral equations was derived by Petcharat *et al.* (2012) when observations are the first order of moving average process, MA(1), with exponential white noise. Phanyaem *et al.* (2014) studied analytical exact formulas of ARL_0 and ARL_1 using integral equation and NIE method for CUSUM control chart for ARMA(1,1) process with exponential distribution white noise. Recently, Petcharat *et at.* (2015) derived the explicit formulas of ARLs for CUSUM control chart when observations are the q order moving average, MA(q), with exponential white noise using the integral equation. The integral equation was based on Fredholm integral equations of the second kind. Finally, Peerajit *et al.* (2016) presented the numerical integral equation (NIE) method of ARLs on CUSUM control chart for long memory process with an ARFIMA model with exponential white noise.

The aim of this paper is to present the explicit formulas and numerical integral equation (NIE) method for ARFIMA process. In section 2, the long memory process for ARFIMA model on CUSUM control chart is presented. In section 3, the uniqueness of solution by using Banach fixed point theorem is described (Venkateshwara *et al.*, 2001). In section 4, the solutions of the integral equation for ARLs are presented and the comparison of analytical results between explicit formulas and NIE method is presented in section 5. Finally, section 6 summarizes the real applications in this paper along with a few topics for further research.

2. The Long Memory Process for ARFIMA Model on CUSUM Control Chart

The CUSUM control chart was the first introduced by Page (1954) to detect small shifts in the mean of a process and is now widely implemented in process control. Let ξ_t be observations of a stationary autoregressive fractionally integrated moving average (ARFIMA) process of order (p,d,q), denoted by ARFIMA(p,d,q) with exponential white noise. The ARFIMA(p,d,q) process shows the characteristic of long memory when the parameter d (the degree of differencing) takes values between (0, 0.5) (Baillie, 1996; Granger & Joyeux, 1980; Hosking, 1981).

The general form of the ARFIMA(p,d,q) process (X_t) which is used on CUSUM control charts has the following form:

$$X_t =$$
$$\mu + \xi_t - \theta_1\xi_{t-1} - \theta_2\xi_{t-2} - \dots - \theta_q\xi_{t-q}$$
$$-\left(-dX_{t-1} + \frac{d(d-1)}{2!}X_{t-2} - \frac{d(d-1)(d-2)}{3!}X_{t-3} + \dots\right)$$
$$+\left(\phi_1 X_{t-1} - d\phi_1 X_{t-2} + \frac{d(d-1)}{2!}\phi_1 X_{t-3} - \frac{d(d-1)(d-2)}{3!}\phi_1 X_{t-4} + \dots\dots\right)$$
$$+\left(\phi_2 X_{t-2} - d\phi_2 X_{t-3} + \frac{d(d-1)}{2!}\phi_2 X_{t-4} - \frac{d(d-1)(d-2)}{3!}\phi_2 X_{t-5} + \dots\right)$$
$$\vdots$$
$$+\left(\phi_p X_{t-p} - d\phi_p X_{t-p-1} + \frac{d(d-1)}{2!}\phi_p X_{t-p-2} - \frac{d(d-1)(d-2)}{3!}\phi_p X_{t-p-3} + \dots\right), \quad (1)$$

where ξ_t is a white noise process assumed with exponential distribution ($\xi_t \sim Exp(\alpha)$). The initial value is normally the process mean, $|\phi_i| < 1$ is an autoregressive coefficient; $i = 1,2,\dots,p$ and $|\theta_i| < 1$ is a moving average coefficient; $i = 1,2,\dots,q$. It is assumed the initial value of (ARFIMA (p,d,q) processes $\xi_{t-1}, \xi_{t-2}, \dots, \xi_{t-q}, X_{t-1}, \dots, X_{t-p}, X_{t-p-1}, \dots$ equal 1 and μ is a constant.

The CUSUM chart based on ARFIMA(p,d,q) process is defined by the following recursion:

$$Y_t = \max(Y_{t-1} + X_t - a,\ 0), \quad t = 1,2,\dots,\ Y_0 = u, \quad (2)$$

where Y_t is the CUSUM statistic, X_t is a sequence of ARFIMA (p,d,q) process, the starting value $Y_0 = u$, u is an initial value and a is a reference value of CUSUM chart.

The corresponding stopping time (τ_b) for (2) is defined as:

$$\tau_b = \inf\{t > 0;\ Y_t > b\},\ u < b, \quad (3)$$

where b is a constant on known parameter as the Upper Control Limit (UCL).

3. Uniqueness of Solution of Integral Equation for the ARLs

The ARLs of the CUSUM control chart are defined as $C(u) = \mathbb{E}_u(\tau_b)$. The notation $\mathbb{P}_y$ denotes the probability measure, the notation $\mathbb{E}_y$ denotes the induced expectation corresponding to the initial value $Y_0 = u$, and $C(u)$ denotes the ARLs of ARFIMA process on CUSUM chart. Then the function of $C(u)$ is initial value $u; u \in [0,b]$, which can be shown by the ARLs (Mititelu et al., 2010; Venkateshwara et al., 2001), defined as $\text{ARL} = C(u) = \mathbb{E}_u(\tau_b) < \infty$, is the unique of solution of integral equation for ARLs as follows:

$$C(u) = 1 + \mathbb{E}_y[I\{0 < Y_1 < b\}C(Y_1)] + \mathbb{P}_y\{Y_1 = 0\}C(0), \quad (4)$$

where

$$I(0 < Y_1 < b) = \begin{cases} 1 & ;\ 0 < Y_1 < b \\ 0 & ;\ Otherwise \end{cases}$$

is the indicator function.

Let ξ_t is continuous distribution i.i.d random variable with exponential distributed given by $F(u) = 1 - e^{-\alpha u}$ and $f(u) = \frac{dF(u)}{du} = \alpha e^{-\alpha u}$ have been proposed in (Mititelu et al., 2010; Mititelu et al., 2011). Hence, the integral equation of ARFIMA(p,d,q) process on CUSUM control chart can be written in the below form:

$$C(u) = 1 + \alpha e^{\alpha(u-a+X_t)}\int_0^b C(z)e^{-\alpha z}dz$$
$$+\left[1 - e^{-\alpha(a-u-X_t)}\right]C(0),\ u \in [0,a). \quad (5)$$

Obviously, the right-hand side of the equation (5) becomes a continuous function, so the solutions of the integral equation (5) is also a continuous function.

Theorem 3.1 (Banach fixed point theorem)

Let $H(I)$ be a non-empty and closed set in a Banach space. Assume that $T: H(I) \to H(I)$ is a contraction mapping, with contraction constant $q \in [0,1)$, i.e.,

$$\|T(C_1) - T(C_2)\| \le q\|C_1 - C_1\|\ ;\ \forall C_1, C_1 \in H(I).$$

Then there exists a unique $C(.) \in I$ such that $T(C(u)) = C(u)$, i.e. T has a unique fixed point in (Sofonea et al., 2006).

Now, consider the non-empty and closed set in a Banach space $(H(I), \|\ \|_\infty)$, where $H(I)$ is the space of all continuous functions on a compact interval I; $I = [0,\ a]$ and $\|\ \|_\infty$ is the sup norm defined as $\|C\|_\infty = \sup_{u\in I}|C(u)|$. This norm is also called the supremum norm for all $u \in [0,\ a]$ and $C(.) \in H(I)$ (Venkateshwara et al., 2001). In this case, let T be an operator in the class of all continuous functions $H(I)$ where I is a compact interval; $I = [0,\ a]$ and define the operators T by

$$T(C(u)) =$$
$$1 + \alpha e^{\alpha(u-a+X_t)}\int_0^b C(z)e^{-\alpha z}dz + (1 - e^{-\alpha(a-u-X_t)})C(0). \quad (6)$$

Therefore, the operator T in (6) can be map $H(I)$ into $H(I)$. The following well-known of the Banach fixed point theorem, if the operator T is a contraction, then the fixed point equations $T(C(u)) = C(u)$ have a unique solution (Venkateshwara et al., 2001). To prove the uniqueness of the solution of (6) the following theorem in 3.2 is considered.

Theorem 3.2 The operator T is the contraction on a metric space $(H(I), \|\ \|_\infty)$ with the norm $\|C\|_\infty = \sup_{u\in I}|C(u)|$.

Proof: To show that T is the contraction and to prove that for all $u \in I$, and two arbitrary function $C_1, C_2 \in H(I)$ in According to (6) one should achieve the following

$$\|T(C_1)-T(C_2)\|_\infty$$

$$= \sup_{u\in I}|C(u)|$$

$$\le \sup_{u\in I}\left\{\left|\left(C_1(0)-C_2(0)\right)\left(1-e^{-\alpha(a-u-X_t)}\right) + \alpha e^{\alpha(u-a+X_t)}\int_0^b \left(C_1(z)-C_2(z)\right)e^{-\alpha z}dz\right|\right\}$$

By Triangular inequality

$$|C_1(0)-C_2(0)| \le \sup_{u\in I}|C_1(u)-C_2(u)| = \|C_1-C_2\|_\infty .$$

$$= \|C_1-C_2\|_\infty \sup_{u\in I}\left\{\left(1-e^{-\alpha(a-u-X_t)}\right)+\alpha e^{\alpha(u-a+X_t)}\int_0^b e^{-\alpha z}dz\right\}$$

$$= \sup_{u\in I}\left\{1-e^{-\alpha(a-u-X_t)-\alpha b}\right\}\|C_1-C_2\|_\infty$$

$$= \left(1-e^{\alpha X_t-\alpha b}\right)\|C_1-C_2\|_\infty .$$

Thus, $\|T(C_1)-T(C_2)\|_\infty = q\|C_1-C_2\|_\infty$,

where $q=\left[1-e^{\alpha X_t-\alpha b}\right]\in[0,1)$, and q is a positive constant.

Therefore, T is the contraction mapping in the non-empty and closed set in a Banach space, with contraction constant $q\in[0,1)$, then there exist a unique of solution such that $T\left(C(u)\right)=C(u)$. By Theorem 3.2 and Banach fixed point theorem which was used to guarantee the existence and uniqueness of the solution for ARL.

4. The Solutions of Integral Equation for ARLs

4.1 The explicit formulas

The derived explicit formulas from the solution of integral equation (5) for ARLs are presented as follows:

Theorem 4.1 The solutions of $T\left(C(u)\right)=C(u)$ is

$$C(u)=e^{\alpha b}(1+e^{\alpha(a-X_t)}-\alpha b)-e^{\alpha u},\ u\in[0,a]. \quad (7)$$

Proof: According to (5), we have that

$$C(u)= 1+\alpha e^{\alpha(u-a+X_t)}\int_0^b C(z)e^{-\alpha z}dz+(1-e^{-\alpha(a-u-X_t)})C(0),$$
$$u\in[0,a].$$

Let $s=\int_0^b C(z)e^{-\alpha z}dz.$ The function $C(u)$ can be rewritten as

$$C(u)=1+\alpha e^{\alpha(u-a+X_t)}s+(1-e^{-\alpha(a-u-X_t)})C(0). \quad (8)$$

In particular $u=0$, if, then

$$C(0)=1+\alpha e^{\alpha(u-a+X_t)}s+(1-e^{-\alpha(a-u-X_t)})C(0),\ \text{Thus}$$

$$C(0)=e^{\alpha(a-X_t)}+\alpha s. \quad (9)$$

Substituting (9) into (8) then $C(u)$ as formed

$$C(u)= 1+\alpha e^{\alpha(u-a+X_t)}s+(1-e^{-\alpha(a-u-X_t)})(e^{\alpha(a-X_t)}+\alpha s).$$

Consequently,

$$C(u)=1+\alpha s+e^{\alpha(a-x_t)}-e^{\alpha u}. \quad (10)$$

Finding a constant s from (10) as formed

$$s=\int_0^b(1+\alpha s+e^{\alpha(a-X_t)}-e^{\alpha y})e^{-\alpha y}dy$$
$$=\int_0^b(1+\alpha s+e^{\alpha(a-X_t)})e^{-\alpha y}dy-\int_0^b e^{\alpha y}e^{-\alpha y}dy.$$

Here, a constant s can be rewritten

$$s=\frac{e^{\alpha b}}{\alpha}(1-e^{-\alpha b})(1+e^{\alpha(a-X_t)})-be^{\alpha b}. \quad (11)$$

Finally, substituting a constant s into (10)

$$C(u)$$
$$=1+\alpha\left(\frac{e^{\alpha b}}{\alpha}(1-e^{-\alpha b})(1+e^{\alpha(a-X_t)})-be^{\alpha b}\right)+e^{\alpha(a-X_t)}-e^{\alpha u}$$
$$=1+e^{\alpha b}(1-e^{-\alpha b})(1+e^{\alpha(a-X_t)})-\alpha be^{\alpha b}+e^{\alpha(a-X_t)}-e^{\alpha u}$$
$$=e^{\alpha b}(1+e^{\alpha(a-X_t)}-\alpha b)-e^{\alpha u} \qquad ;u\ge 0.$$

Therefore, the explicit formulas for ARL_0 and ARL_1 on CUSUM chart can be written:

$$ARL_0=e^{\alpha_0 b}(1+e^{\alpha_0(a-X_t)}-\alpha_0 b)-e^{\alpha_0 u}, \quad (12)$$

and

$$ARL_1=e^{\alpha_1 b}(1+e^{\alpha_1(a-X_t)}-\alpha_1 b)-e^{\alpha_1 u}. \quad (13)$$

4.2 Numerical Integral Equation (NIE) Method

This section the authors presents the numerical integral equation (NIE) method to compute the solutions $C(u)=\mathbb{E}_u(\tau_b)<\infty$ of integral equations (5) to extend the function $C(u)$ into the Fredholm integral equations of the second kind (Wieringa, 1999) as the following form:

$$C(u)=1+\int_0^b C(z)f(z+a-u-X_t)dz+C(0)F(a-u-X_t), \quad (14)$$

where $F(u)=1-e^{-\alpha u}$ and $f(u)=\frac{dF(u)}{du}=\alpha e^{-\alpha u}$.

Let $\tilde{C}(u)$ denote the approximated numerical integral equation (14) using the Gauss-Legendre quadrature rule as follows:

$$\tilde{C}(a_i) \approx 1 + \sum_{j=1}^{m} w_j \tilde{C}(a_j) f(a_j + a - a_i - X_t) + \tilde{C}(0) F(a - a_i - X_t), \quad (15)$$

where

$i = 1, 2, ..., m,$ w_j is a weight define different quadrature rules with $w_j = \frac{b}{m} \ge 0,$

a_j is a set of point with $a_j = \frac{b}{m}\left(j - \frac{1}{2}\right);\ j = 1, 2, ..., m.$

The previous equation is a system of m linear equations in the m unknowns $\tilde{C}(a_1), \tilde{C}(a_2), ..., \tilde{C}(a_m)$, which can be rearranged as:

$$\tilde{C}(a_1) = 1 + \sum_{j=2}^{m} w_j \tilde{C}(a_j) f(a_j + a - a_1 - X_t) + \tilde{C}(a_1)\left[F(a - a_1 - X_t) + w_1 f(a - X_t)\right]$$

$$\tilde{C}(a_2) = 1 + \sum_{j=2}^{m} w_j \tilde{C}(a_j) f(a_j + a - a_2 - X_t) + \tilde{C}(a_1)\left[F(a - a_2 - X_t) + w_1 f(a_1 + a - a_2 - X_t)\right]$$

$$\vdots$$

$$\tilde{C}(a_m) = 1 + \sum_{j=2}^{m} w_j \tilde{C}(a_j) f(a_j + a - a_m - X_t) + \tilde{C}(a_1)\left[F(a - a_m - X_t) + w_1 f(a_1 + a - a_m - X_t)\right].$$

It can be written in matrix form as:

$$\mathbf{C}_{m\times 1} = \mathbf{1}_{m\times 1} + \mathbf{R}_{m\times m}\mathbf{C}_{m\times 1}, \quad (16)$$

where $\mathbf{C}_{m\times 1} = \begin{bmatrix} \tilde{C}(a_1) \\ \tilde{C}(a_2) \\ \vdots \\ \tilde{C}(a_m) \end{bmatrix}, \quad \mathbf{1}_{m\times 1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix},$

$$\mathbf{R}_{m\times m} = \begin{bmatrix} F(a - a_1 - X_t) + w_1 f(a - X_t) & \ldots & w_m f(a_m + a - a_1 - X_t) \\ F(a - a_1 - X_t) + w_1 f(a_1 + a - a_2 - X_t) & \ldots & w_m f(a_m + a - a_2 - X_t) \\ \vdots & & \vdots \\ F(a - a_m - X_t) + w_1 f(a_1 + a - a_m - X_t) & \ldots & w_m f(a_m + a - a_m - X_t) \end{bmatrix},$$

and $I_m = diag(1, 1, \ldots, 1)$ is identity matrix order m.

Therefore, $\mathbf{C}_{m\times 1} = \mathbf{1}_{m\times 1} + \mathbf{R}_{m\times m}\mathbf{C}_{m\times 1}$, or similarly $(\mathbf{1}_m - \mathbf{R}_{m\times m})\mathbf{C}_{m\times 1} = \mathbf{1}_{m\times 1}$. If $(1 - \mathbf{R}_{m\times m})$ is invertible and exists, then the unique solution of matrix equation (16) is achieved as follows:

$$\mathbf{C}_{m\times 1} = (\mathbf{1} - \mathbf{R}_{m\times m})^{-1}\mathbf{1}_{m\times 1}. \quad (17)$$

Consequently, the numerical integral equation (NIE) method for ARL on CUSUM chart can be written

$$\tilde{C}(u) \approx 1 + \sum_{j=1}^{m} w_j \tilde{C}(a_j) f(a_j + a - u - X_t) + \tilde{C}(0) F(a - u - X_t), \quad (18)$$

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

5. Comparison of Analytical Results

This section derives the explicit formulas and NIE method values for ARL_0 and ARL_1 from equation (12), (13), and (18) with parameters a and b for ARL_0 which fixes at 370 and 500. Also, the explicit formulas values are compared with values obtained from the NIE method under these same parameters.

The comparison of efficiency based on percentage of absolute difference (*Diff* (%)) is defined as:

$$Diff\,(\%) = \frac{|\,C(u) - \tilde{C}(u)\,|}{C(u)} \times 100\%. \quad (19)$$

where $C(u)$ is ARLs of the explicit formulas values, and $\tilde{C}(u)$ is ARLs of the NIE method values.

Criteria for consideration; If the *Diff* (%) is less than 2%, then ARL values from the explicit formulas and the NIE method are similar and in good agreement.

From the results in Tables 1-4, the authors applied equation (12), (13), and (18) to evaluate the ARLs for the long memory process on ARFIMA (2,0.2,1) model. The comparison of efficiency between the explicit formulas and NIE method with given a = 3, 3.5, ϕ_1 = 0.10, -0.10, ϕ_2 = 0.20, and θ_1 = 0.10 for ARL_0 = 370 and 500 are shown.

The process in control parameter value (α_0) with shift size (δ = 0) had a fixed ARL_0 = 370 and 500. The first row in Tables 1-4 shows that the values of ARL_0 in explicit formulas were close to the NIE method and also approached 370 and 500. The computational CPU time of ARLs by NIE method was computed. The values in parentheses represent the CPU time for calculation with division points, m = 800 nodes. The CPU time with the NIE method was about 1.8-1.9 hours, this was very high compared to the explicit formulas which equaled less than 1 second.

On the other hand, the process out-of-control was presented with parameter values, $\alpha_1 = \alpha_0(1 + \delta)$ where δ = 0.01, 0.03, 0.05, 0.10, 0.20, and 0.40. According to the results from Tables 1-4, the percentage of absolute difference of the explicit formulas and NIE method was less than 0.23% calculated using equation (19). In summary, the CPU time of the explicit formulas was less than one second, while the NIE method was approximately at 1.8-1.9 hours.

Table 1. Comparison of ARL values for ARFIMA(2, 0.2, 1) process using explicit formulas against NIE method when given $a = 3$, $\phi_1 = 0.10$, $\phi_2 = 0.20$ and $\theta_1 = 0.10$, $b = 3.56928$ for $ARL_0 = 370$ and $b = 3.900538$ for $ARL_0 = 500$.

Shift size (δ)	$ARL_0 = 370$			$ARL_0 = 500$		
	explicit formulas	NIE	Diff(%)	explicit formulas	NIE	Diff(%)
0.00	370.0004	369.2444 (1.97)	0.2043	500.0004	498.8793 (1.83)	0.2242
0.01	347.0438	346.3443 (1.98)	0.2016	467.1991	466.1663 (1.85)	0.2211
0.03	306.4437	305.8424 (1.84)	0.1962	409.5039	408.6236 (1.86)	0.2150
0.05	271.8672	271.3478 (1.84)	0.1910	360.7195	359.9651 (1.86)	0.2091
0.10	205.4008	205.0331 (1.83)	0.1790	268.0102	267.4862 (1.86)	0.1955
0.20	125.7785	125.5799 (1.85)	0.1579	159.3739	159.1001 (1.85)	0.1718
0.40	58.4003	58.3274 (1.82)	0.1248	70.6982	70.6029 (1.86)	0.1348

The values in parentheses are CPU times in numerical integration Equation methods (Hours)

Table 2. Comparison of ARL values for ARFIMA(2, 0.2, 1) process using explicit formulas against NIE method when given $a = 3.5$, $\phi_1 = 0.10$, $\phi_2 = 0.20$ and $\theta_1 = 0.10$, $b = 2.9450131$ for $ARL_0 = 370$ and $b = 3.2604379$ for $ARL_0 = 500$.

Shift size (δ)	$ARL_0 = 370$			$ARL_0 = 500$		
	explicit formulas	NIE	Diff(%)	explicit formulas	NIE	Diff(%)
0.00	370.0003	369.3511 (1.83)	0.1755	499.9996	499.0261 (1.82)	0.1947
0.01	347.9397	347.3364 (1.86)	0.1734	468.5892	467.6878 (1.85)	0.1924
0.03	308.7699	308.2471 (1.92)	0.1693	413.0919	412.3162 (1.85)	0.1878
0.05	275.2379	274.7825 (1.83)	0.1655	365.8878	365.2167 (1.85)	0.1834
0.10	210.2388	209.9102 (1.84)	0.1563	275.3254	274.8488 (1.85)	0.1731
0.20	131.1087	130.9250 (1.85)	0.1401	167.2381	166.9789 (1.85)	0.1550
0.40	62.3716	62.3003 (1.82)	0.1143	76.3363	76.2399 (1.86)	0.1263

The values in parentheses are CPU times in numerical integration Equation methods (Hours)

Table 3. Comparison of ARL values for ARFIMA(2, 0.2, 1) process using explicit formulas against NIE method when given $a = 3$, $\phi_1 = -0.10$, $\phi_2 = 0.20$ and $\theta_1 = 0.10$, $b = 3.390216$ for $ARL_0 = 370$ and $b = 3.715676$ for $ARL_0 = 500$.

Shift size (δ)	$ARL_0 = 370$			$ARL_0 = 500$		
	explicit formulas	NIE	Diff(%)	explicit formulas	NIE	Diff(%)
0.00	369.9995	369.2712 (1.83)	0.1968	499.9999	498.9174 (1.84)	0.2165
0.01	347.3501	346.6752 (1.83)	0.1943	467.6790	466.6798 (1.84)	0.2137
0.03	307.2381	306.6563 (1.82)	0.1894	410.7392	409.8846 (1.85)	0.2081
0.05	273.0155	272.5115 (1.85)	0.1846	362.4938	361.7591 (1.85)	0.2027
0.10	207.0386	206.6794 (1.85)	0.1735	270.5047	269.9903 (1.86)	0.1902
0.20	127.5615	127.3651 (1.84)	0.1540	162.0226	161.7499 (1.87)	0.1683
0.40	59.7016	59.6281 (1.85)	0.1231	72.5572	72.4600 (1.85)	0.1340

The values in parentheses are CPU times in numerical integration Equation methods (Hours)

Table 4. Comparison of ARL values for ARFIMA(2, 0.2, 1) process using explicit formulas against NIE method when given $a = 3.5$, $\phi_1 = -0.10$, $\phi_2 = 0.20$ and $\theta_1 = 0.10$, $b = 2.791475$ for $ARL_0 = 370$ and $b = 3.1044675$ for $ARL_0 = 500$.

Shift size (δ)	$ARL_0 = 370$			$ARL_0 = 500$		
	explicit formulas	NIE	Diff(%)	explicit formulas	NIE	Diff(%)
0.00	370.0004	369.3810 (1.84)	0.1674	499.9999	499.0672 (1.83)	0.1865
0.01	348.0942	347.5182 (1.85))	0.1655	468.8270	467.9628 (1.85)	0.1843
0.03	309.1733	308.6734 (1.85)	0.1617	413.7088	412.9638 (1.85)	0.1801
0.05	275.8254	275.3895 (1.85)	0.1580	366.7811	366.1357 (1.86)	0.1760
0.10	211.0932	210.7776 (1.85)	0.1495	276.6068	276.1467(1.85)	0.1663
0.20	132.0726	131.8952 (1.85)	0.1343	168.6498	168.3979 (1.86)	0.1494
0.40	63.1192	63.0496 (1.85)	0.1103	77.3908	77.2961 (1.85)	0.1224

The values in parentheses are CPU times in numerical integration Equation methods (Hours)

6. Conclusions

This paper presented the investigation of explicit formulas for the average run lengths of long memory process with the ARFIMA(p,d,q) on CUSUM control charts with exponential white noise. The accuracy of the proposed explicit formulas in terms of percentage of absolute difference of the explicit formulas and NIE method were checked and compared. The results showed that both methods were similar and in good agreement with the percentage of absolute difference at less than 0.23%. But, the computational CPU time of the explicit formula was less than one second, while the NIE method was approximately 1.8-1.9 hours. Therefore, the explicit formulas are a preferred alternative to the NIE method because ARL values use a drastically lower computational CPU time.

In conclusion, from the above results, one can see that the explicit formulas and numerical integral equation (NIE) method of ARFIMA(p,d,q) process with exponential white noise on CUSUM control chart can be successfully applied to real world applications for different processes of data, for example in economics, agriculture.

Acknowledgements

This work was supported in part by the Graduate College, King Mongkut's University of Technology North Bangkok. The authors are highly grateful to the referees for their constructive comments and suggestions which helped to improve this research.

References

Areepong, Y. (2009). *An integral equation approach for analysis of control charts* (Doctoral thesis, University of Technology, Sydney, Australia).

Baillie, R. T. (1996). Long memory processes and fractional integration in econometrics. *Journal of Econometrics*, *73*, 5-59.

Barnard, G. A. (1954). Sampling inspection and statistical decisions. *Journal of the Royal Statistical Society: Series B*, *16*(2), 151-174.

Beran, J. (1994). *Statistics for long memory processes*. New York, NY: Chapman and Hall.

Bissell, A. F. (1969). CUSUN techniques for quality control. *Applied Statistics*, *18*, 1-30.

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

Busaba, J., Sukparungsee, S., & Areepong, Y. (2012). Analytical of ARL for trend stationary first order of autoregressive observations on CUSUM procedure. *Proceedings of the World Congress on Engineering*, 103-108.

Caballero, R., Jewson, S., & Brix, A. (2002). Long memory in surface air temperature: detection, modeling, and application to weather derivative valuation. *Climate Research*, *21*, 127-140.

Champ, C. W., & Woodall, W. H. (1987). Exact results for Shewhart control charts with supplementary run rules. *Technometrics*, *29*, 393-399.

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.

Chung, K. J. (1992). Economically optimal determination of the parameters of CUSUM charts (cumulative sum control chart). *International Journal of Quality and Reliability Management*, *9*, 8-17.

Ewan, W. D., & Kemp, K. W. (1960). Sampling inspection of continuous processes with no autocorrelation between successive results. *Biometrika*, *47*, 363-380.

Ewan, W. D. (1963). When and How to use CUSUM charts. *Technometrics*, *11*, 307-315.

Fellner, W. H. (1990). Algorithm as 258. Average run length for cumulative sum schemes. *Applied Statistics*, *39*, 402-412.

Gan, F. F. (1992). Exact run length distribution for one-sided exponential CUSUM schemes. *Statistica Sinica*, *2*, 297-312.

Gan, F. F. (1996). Average run lengths for cumulative sum control chart under liner trend. *Applied Statistics*, *45*, 505-512.

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

Goel, A. L., & Wu, S. M. (1971). Determination of A.R.L. and a contour nomogram for CUSUM charts to control normal mean. *Technometrics*, *13*, 221-230.

Hawkins, D. M. (1992). A fast accurate approximation for average run lengths of CUSUM control charts. *Journal of Quality Technology*, *24*, 37-43.

Hawkins, D. M., & Olwell, D. H. (1998). Cumulative sum charts and charting for quality improvement. New York, NY: Springer.

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

Jacob, P. A., & Lewis, P. A. W. (1977). A mixed autoregressive-moving average exponential sequence and point process (EARMA 1,1). *Advances in Applied Probability*, *9*(1), 87-104.

Johnson, N. L., & Leone, F. C. (1962). Cumulative sum control charts: Mathematical principles applied to their construction and use. *Industrial Quality Control*, *19*, 22-28.

Lashkari, R. S., & Rahim, M. A. (1982). An Economic design of cumulative sum charts to control non-normal process means. *Computers and Industrial Engineering*, *6*, 1-18.

Lucas, J. M. (1976). The design and use of cumulative sum control schemes. *Technometrics*, *14*, 51-59.

Luceño, A., & Puig-Pey, J. (2000). Evaluation of the run length probability distribution for CUSUM charts: Assessing chart performance. *Technometrics*, *42*, 411-416.

Luceño, A., & Puig-Pey, J. (2002). Computing the run length probability distribution for CUSUM charts. *Journal of Quality Technology*, *34*(2), 209-215.

Mititelu, G., Areepong, Y., Sukparungsee, S., & Novikov, A. (2010). Explicit analytical solutions for the average run length of CUSUM and EWMA charts, *East West Journal of Mathematics*, *1*, 253-265.

Mititelu, G., Mazalov, V. V., & Novikov, A. (2011). On CUSUM procedure for hyper-exponential distribution. *Statistics and Probability Letters*.

Mohamed, I., & Hocine, F. (2010). Bayesian estimation of an AR(1) process with exponential white noise. *A Journal of Theoretical and Applied Statistics*, *37* (5), 365-372.

Page, E. S. (1954). Continuous inspection schemes. *Biometrika*, *41*(1-2), 100-114.

Palma, W. (2007). *Long-memory time series: theory and methods*. New York, NY: Wiley.

Pan, J. N., & Chen, S. T. (2008). Monitoring long-memory air quality data using ARFIMA model. *Environmetrics*, *19*, 209-219.

Peerajit, W., Areepong, Y., & Sukparungsee, S. (2016). Numerical integral equation method of average run length of cumulative sum control chart for long memory process with ARFIMA model. *Proceedings of the International MultiConference of Engineers and Computer Scientists* (IMECS), 852-855.

Petcharat, K., Areepong, Y., Sukparungsri, S., & Mititelu, G. (2012). Numerical approximation to the performance of CUSUM charts for EMA (1) process. *International Journal of Mathematical and Computational Sciences*, *6*(12), 1770-1774.

Petcharat, K., Sukparungsee, S., & Areepong, Y. (2015). Exact solution of the average run length for the cumulative sum chart for a moving average process of order *q*. *Science Asia*, *41*, 141-147.

Phanyaem, S., Areepong, Y., Sukparungsee, S., & Mititelu, G. (2014). Explicit formulas of average run length for ARMA(1, 1) process of CUSUM control chart. *Far East Journal of Mathematical Sciences*, *90*(2), 211-224.

Proietti, T. (2014). Exponential smoothing, long memory and volatility prediction. *Centre for Economic and International studies Tor Vergata*, *12*(7), 1-28.

Rabyk, L., & Schmid, W. (2016). EWMA control charts for detecting changes in the mean of a long-memory process, *Metrika*, *79*, 267-301. doi: 10.1007/s00184-015-0555-7

Ramjee, R. (2000). *Quality control charts and persistent processes* (Doctoral thesis, Stevens Institute of Technology, Hoboken, New Jersey, U.S.A.).

Ramjee, R., Crato, N., & Ray, B. K. (2002). A note on moving average forecasts of long memory processes with an application to quality control. *International Journal of Forecasting*, *18*, 291-297.

Ryan, T. P. (1989). *Statistical methods for quality improvement*. New York, NY: John Wiley and Sons.

Sofonea, M., Han, W., & Shillor, M. (2006). *Analysis and approximation of contact problems with adhesion or damage*, New York, NY: Chapman and Hall/CRC Pure and Applied Mathematics.

Vance, L. C. (1986). Average run lengths of cumulative sum control charts for controlling normal means. *Journal of Quality Technology*, *18*, 189-193.

Venkateshwara, B. R., Ralph, L. D., & Joseph, J. P. (2001). Uniqueness and convergence of solutions to average run length integral equations for cumulative sums and other control charts. *IIE Transactions*, *33*, 463-469.

Wieringa, J. E. *Statistical process control for serially correlated data* (Doctoral thesis, University of Groningen, Groningen, The Netherlands).

Woodall, W. H. (1983). The distribution of the run length of one-sided CUSUM procedures for continuous random variables. *Technometrics*, *24*, 295-301.