

Copula modelling for multivariate statistical process control: a review

Piyapatr Busababodhin^{1,a}, Pimpan Amphanthong^b

^aDepartment of Mathematics, Mahasarakham University, Thailand;

^bDepartment of Mathematics, Rajamangala University of Technology Suvarnabhumi, Thailand

Abstract

Modern processes often monitor more than one quality characteristic that are referred to as multivariate statistical process control (MSPC) procedures. The MSPC is the most rapidly developing sector of statistical process control and increases interest in the simultaneous inspection of several related quality characteristics. Most multivariate detection procedures based on a multi-normality assumptions are independent, but there are many processes that assume non-normality and correlation. Many multivariate control charts have a lack of related joint distribution. Copulas are tool to construct multivariate modelling and formalizing the dependence structure between random variables and applied in several fields. From copula literature review, there are a few copula to apply in MSPC that have multivariate control charts, and represent a successful tool to identify an out-of-control process. This paper presents various types of copulas modelling for the multivariate control chart. The performance measures of the control chart are the average run length (ARL) and the average number of observations to signal (ANOS). Furthermore, a Monte Carlo simulation is shown when the observations were from an exponential distribution.

Keywords: copula, Spearman's rho, multivariate control chart, copula Markov CUSUM, measuring dependence, performance characteristics, average run length

1. Introduction

The previously mentioned quality control charts deal with controlling one characteristic and are applicable for processes that have one variable or more, which are referred to as univariate or multivariate control charts, respectively. In real-world applications, especially manufacturing industries, the existing correlation between characteristics makes it more coherent to monitor two or more quality characteristics simultaneously rather than single ones separately. In this regards, if the corresponding quality characteristics are all measured in a nominal or ordinal scale, the related control charts are named “multivariate attributes charts”. However, if the characteristics are measured in an interval or ratio scale, the corresponding control charts are entitled “multivariate variables charts”, both of them are called “multivariate control charts”.

Multivariate control charts are used to simultaneously monitor several quality characteristics to detect the mean changes in manufacturing industries. They are a powerful tool in multivariate statistical process control (MSPC) to identify an out-of-control process. Generally, multivariate detection procedures are based on multi-normality assumption and independence; however, many processes

¹ Corresponding author: Department of Mathematics, Mahasarakham University, Mahasarakham, 44150, Thailand.
E-mail: piyapatr.b@msu.ac.th

exhibit non-normality and correlation. Most multivariate control charts are generalizations of the univariate counterparts (Mohmoud and Maravelakis, 2013), such as the multivariate Shewhart control chart, the multivariate exponentially weighted moving average (MEWMA) control chart proposed by Lowry *et al.* (1992), and the multivariate cumulative sum (MCUSUM) control chart (Bersimis *et al.*, 2005, 2007). Many practitioners have addressed the problem of correlated data in MSPC (Lowry and Montgomery, 1995), and multivariate control charts with the lack of related joint distribution.

Recently, copulas modelling is a general approach to model multivariate non-normal data (Joe, 1997), with the dependence structure separated from the univariate margins. Copulas create a link between multivariate joint distributions and univariate marginal distributions; they have been widely studied and applied in areas concerned with the problem of dependence relations. Many researchers have developed the copula for use with control charts (Dokouhaki and Noorossana, 2013; Fatahi *et al.*, 2011, 2012; Hryniewicz, 2012; Hryniewicz and Szediw, 2010; Kuvattana *et al.*, 2016; Verdier, 2013). Time is used to represent some attributes or variable measures in the manufacturing processes.

Several multivariate control charts have been proposed (Bourke, 1991; El-Midany *et al.*, 2010; Montgomery, 2013; Xie and Goh, 1992) and presented the copula approach that use to make the joint distribution of two data series; however, few investigations have been reported on copula modelling for a multivariate control chart. This paper is organized as: Section 2 contains some MSPC. In Section 3 will introduce the basic concepts of copulas and the dependence structure by calculating Spearman's rho and Kendall's tau. Section 4 presents the application of copula to multivariate control chart. The basic characteristics that describes the performance of control charts will be introduced in Section 5. Finally, the discussion and conclusion are presented in Section 6.

2. Multivariate statistical process control

In practice, there are many situations in which the simultaneous monitoring of two or more related quality characteristics is necessary. Process monitoring in several related variables are collectively known as MSPC. The most useful tool of MSPC is the quality control chart (Bersimis *et al.*, 2007), of which the multivariate control chart is one of the main practice tools in multivariate quality control. This section presents some popular and widely used multivariate control charts.

2.1. The multivariate Shewhart control chart

Consider two random variables and suppose that $\bar{\mathbf{X}}$ and \mathbf{S} are the sample mean vector and covariance of matrix, respectively. Let m and p be sample and the number of quality characteristics observed in each sample, respectively. The Hotelling T^2 statistic is

$$T^2 = (\mathbf{X} - \bar{\mathbf{X}})' \mathbf{S}^{-1} (\mathbf{X} - \bar{\mathbf{X}}), \quad (2.1)$$

where

$$\bar{\mathbf{X}} = \frac{1}{m} \sum_{i=1}^m X_i \quad \text{and} \quad \mathbf{S} = \frac{1}{m-1} \sum_{i=1}^m (X_i - \bar{\mathbf{X}})(X_i - \bar{\mathbf{X}})'$$

SPC is usually split into two phases, namely phase I and phase II, as explained in Montgomery (2013). Phase I constitutes a retrospective analysis, constructing trial control limits to determine if the process has been in-control; subsequently, the controlled data are used in phase II to monitor the process. The control chart contains a center line that represent the average value of the quality characteristic corresponding to the in-control process. Two other horizontal line, called the upper control limit (UCL) and the lower control limit (LCL) are shown on the chart.

For the phase I, the control limit are $UCL = \{(m - 1)^2/m\}\beta_{\alpha,p/2,(m-p-1)/2}$ and $LCL = 0$, where $\beta_{\alpha,p/2,(m-p-1)/2}$ is an upper α percentage point of beta distribution with parameter $p/2$ and $(m-p-1)/2$. For the phase II, the control limit for this statistic are $UCL = \{p(m + 1)(m - 1)/(m^2 - mp)\}F_{\alpha,p,(m-p)}$ and $LCL = 0$, where $F_{\alpha,p,(m-p)}$ is F distribution with parameter p and $(m - p)$. This article focuses on phase II control charts and their performance.

2.2. The multivariate cumulative sum control chart

The MCUSUM control chart is the multivariate extension of the univariate cumulative sum (CUSUM) control chart. The MCUSUM chart was initially proposed by Crosier (1988). The MCUSUM control chart may be expressed as:

$$C_i = \left[(S_{t-1} + X_t - a)' \Sigma^{-1} (S_{t-1} + X_t - a) \right]^{\frac{1}{2}}, \quad t = 1, 2, 3, \dots, \tag{2.2}$$

where covariance matrix and the cumulative sums are expressed as

$$S_t = \begin{cases} 0, & \text{if } C_t \leq k, \\ (S_{t-1} + X_t - a) \left(1 - \frac{k}{C_t} \right), & \text{if } C_t > k, \end{cases}$$

where the reference value $k > 0$ and a are the aim point or target value for the mean vector (Khoo *et al.*, 2009). The control chart statistics for MCUSUM chart is

$$Y_t = \left[S_t' \Sigma^{-1} S_t \right]^{\frac{1}{2}}, \quad t = 1, 2, 3, \dots$$

The signal gives an out-of-control if $Y_t > h$, where h is the control limit (Alves *et al.*, 2010).

2.3. The multivariate exponentially weighted moving average control chart

Lowry *et al.* (1992) have developed a MEWMA control chart. The MEWMA control chart is a logical extension of the univariate exponentially weighted moving average (EWMA) control chart. The EWMA statistics assigns less weight to past observations than the current observation (Suriyakat *et al.*, 2012). Suppose that X_i is a $p \times 1$ vector of observation at sample $i = 1, 2, 3, \dots$ with a target mean vector equal to the zero vector and know covariance matrix σ and the vector are independent over time. The extension of the EWMA control chart to the multivariate case is then defined as:

$$Z_i = \lambda X_i + (1 - \lambda) Z_{i-1}, \tag{2.3}$$

where λ ($0 \leq \lambda \leq 1$) is the scalar charting constant which may be adjusted to change the weighting of the past observations and $Z_0 = 0$. The quantity plotted on the control chart of the MEWMA (Montgomery, 2013) is

$$T_i^2 = Z_i \Sigma_{Z_i}^{-1} Z_i.$$

The control chart signals a shift in the mean vector when $T_i^2 > h$, where h is the control limit chosen to achieve a desired in-control and the covariance matrix for Z_i is

$$\Sigma_{Z_i} = \frac{\lambda}{2 - \lambda} \left[1 - (1 - \lambda)^{2i} \right] \Sigma.$$

The average run length (ARL) performance of the MEWMA control chart depends on the off-target mean vector because the process mean vector shifts from the zero vector to a new out-of-control vector. The shift size is reported in terms of a quantity $\delta = (\mu' \Sigma^{-1} \mu)^{1/2}$ which is usually called the non-centrality parameter. The value $\delta = 0$ is the in-control state and large value of δ correspond to bigger shifts in the mean. Note that if $\lambda = 1$, the MEWMA control chart statistic reduce to $T_i^2 = X_i' \Sigma_i^{-1} X_i$, the statistic used for the Hotelling T^2 control chart (Runger *et al.*, 1999).

3. Copula function and measuring dependence

3.1. Copula function

Copulas offer a useful tool in modelling dependence among random variables. Most existing copulas in the literature are symmetric versus real world data that may exhibit an asymmetric nature. The copula function is a multivariate distribution with univariate margins. The multivariate copula function is used to capture the dependence between two or more random variables. Suppose that a random vector has a joint distribution function with continuous marginal distribution function that has uniform distribution, then there exists a unique d -dimensional copula. In this section, the theory that is the central foundation of the copula is described (Genest and McKay, 1986; Joe, 2015; Nelsen, 2016; Sklar, 1959; Trivedi and Zimmer, 2005).

This section presents some definitions and results that are necessary to understand a (bivariate) copula. A copula is a multivariate distribution function defined on \mathbb{I}^n , where $\mathbb{I} := [0, 1]$, with uniformly distributed marginals. This paper focuses on bivariate (two-dimensional, $n = 2$) copulas.

Definition 1. A bivariate copula is a function $C : \mathbb{I}^2 \rightarrow \mathbb{I}$, which satisfies the following properties:

$$(P1) \quad C(0, v) = C(u, 0) = 0, \quad \forall u, v \in \mathbb{I}$$

$$(P2) \quad C(1, u) = C(u, 1) = u, \quad \forall u \in \mathbb{I}$$

$$(P3) \quad C \text{ is 2-increasing, i.e., } \forall u_1, u_2, v_1, v_2 \in \mathbb{I} \text{ with } u_1 \leq u_2, v_1 \leq v_2,$$

$$C(u_2, v_2) + C(u_1, v_1) - C(u_1, v_2) - C(u_2, v_1) \geq 0.$$

The importance of copulas has increased due to their application to several fields of research. Their relevance primarily comes from Sklar's theorem (Sklar, 1959): If X and Y are two continuous random variables with joint distribution function H and marginal distribution functions F and G , respectively, then there exists a unique copula C such that $H(x, y) = C(F(x), G(y))$ for all $(x, y) \in \mathbb{R}^2$. And conversely, given a copula C and two univariate distribution functions F and G , the function H defined above is a joint distribution function with margins F and G . The following Sklar's theorem clarifies the role that copulas play in the relationship between multivariate distribution functions and the univariate margins.

Definition 2. Suppose X and Y are two random variables with marginal distribution functions F and G , respectively. Then Spearman's rho is the ordinary (Pearson) correlation coefficient of the transformed random variables $F(X)$ and $G(Y)$, while Kendall's tau is the difference between the probability of concordance $\Pr[(X_1 - X_2)(Y_1 - Y_2) > 0]$ and the probability of discordance $\Pr[(X_1 - X_2)(Y_1 - Y_2) < 0]$ for two independent pairs (X_1, Y_1) and (X_2, Y_2) of observations drawn from the distribution.

For convenience, we adopt the following notations, for $j = 1, 2, \dots$, where copulas C are defined in Table 1. The list of copulas in Table 1 is considered in our analysis.

Table 1: Nine different copula functions

Copula	Type	Copula function
Normal	Symmetric	$C(u, v) := C_F(u, v)$
Frank	Asymmetric	$C(u, v) := \text{Frank}(u, v, \theta) = -\frac{1}{\theta} \log \left[1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1} \right],$ $\theta \in \mathbb{R} \setminus \{0\}$
Gumbel	Asymmetric	$C(u, v) := \text{Gumbel}(u, v, \theta) = \exp \left[-\left((-\log u)^\theta + (-\log v)^\theta \right)^{\frac{1}{\theta}} \right],$ $\theta \geq 1$
Clayton	Asymmetric	$C(u, v) := \text{Clayton}(u, v, \theta) = \left(u^{-\theta} + v^{-\theta} - 1 \right)^{-\frac{1}{\theta}}, \quad \theta \in (0, \infty)$
AMH	Symmetric	$C(u, v) := \text{AMH}(u, v, \theta) = \frac{uv}{1 - \theta(1-u)(1-v)},$ $\theta \in (-1, 1]$
FGM	Symmetric	$C(u, v) := \text{FGM}(u, v, \theta) = uv + \theta uv(1-u)(1-v), \quad \theta \in (-1, 1]$
Joe	Symmetric	$C(u, v) := \text{Joe}(u, v, \theta) = 1 - \left[(1-u)^\theta + (1-v)^\theta - (1-u)^\theta(1-v)^\theta \right]^{\frac{1}{\theta}},$ $\theta \in [1, \infty)$
Plackett	Symmetric	$C(u, v) := \text{Plackett}(u, v, \theta) = \frac{1 + (\theta - 1)(u + v)}{2(\theta - 1)} - \frac{\sqrt{[1 + (\theta - 1)(u + v)]^2 - 4uv\theta(\theta - 1)}}{2(\theta - 1)},$ $\theta \in [0, \infty)$
Raftery	Symmetric	$C(u, v) := \text{Raftery}(u, v, \theta) = M_2(u, v) + \frac{1 - \theta}{1 + \theta} (uv)^{\frac{1}{1-\theta}} \left[1 - (\max\{u, v\})^{-\frac{1+\theta}{1-\theta}} \right],$ $\theta \in \mathbb{I}$

3.2. Measuring dependence

In terms of dependence properties, Spearman’s rho is a measure of average quadrant dependence; however, Kendall’s tau is a measure of average likelihood ratio dependence. If X and Y are two continuous random variables with copula C , then Spearman’s rho and Kendall’s tau of X and Y are given by,

$$\rho = 12 \iint_{\mathbb{I}^2} C(u, v) \, dudv - 3, \tag{3.1}$$

$$\tau = 4 \iint_{\mathbb{I}^2} C(u, v) \, dC(u, v) - 1. \tag{3.2}$$

Definition 3. A copula C is called absolutely continuous if, when considered as a joint distribution function, $C(u, v)$ has a joint density function given by $c(u, v) := \partial^2 C / (\partial u \partial v)$ and in that case $dC(u, v) = \partial^2 C / (\partial u \partial v) \, dudv$.

Denoting $c(u, v) - 1$ as $h(u, v)$, the following theorem gives a characterization of absolutely continuous copulas.

Theorem 1. A function $C : \mathbb{I}^2 \rightarrow \mathbb{I}$ is an absolutely bivariate copula if and only if there exists a function $h : \mathbb{I}^2 \rightarrow \mathbb{I}$, satisfying the following conditions,

1. Integrability: $\iint_{\mathbb{I}^2} |h(x, y)| \, dx dy < \infty,$
2. Degeneracy: $\int_{\mathbb{I}} h(x, \xi) \, d\xi = \int_{\mathbb{I}} h(\xi, y) \, d\xi = 0 \quad \forall x, y \in \mathbb{I},$

3. *Positive Definiteness:* $h(x, y) \geq -1 \quad \forall (x, y) \in \mathbb{I}^2$, and such that

$$C(u, v) = \int_0^v \int_0^u 1 + h(x, y) dx dy.$$

A copula C is called symmetric if $C(u, v) = C(v, u)$ for all $u, v \in \mathbb{I}$, otherwise asymmetric.

A parametric measure of the linear dependence between random variables is often a correlation coefficient and non-parametric measure of dependence are Spearman's rho and Kendall's tau. Spearman's rho is used to describe the association between random variables with linear dependence. For an association between random variables that are not linear, Kendall's tau is a measurement of the concordance. Kendall's tau is a non-parametric measurement of associations and considered a copula-based dependence measurement. Kendall's tau is straightforward to calculate and used for observation dependencies.

Let X and Y be continuous random variable, then Kendall's tau for X and Y is given by Equation (3.2), where τ_c is Kendall's tau of copula C and the unit square \mathbb{I}^2 is the product $\mathbb{I} \times \mathbb{I}$ where $\mathbb{I} = [0, 1]$. And the expected value of the function $C(u, v)$ of uniform $(0, 1)$ random variables U and V whose joint distribution function is C , i.e.,

$$\tau_c = 4E[C(U, V)] - 1,$$

see Nayland College (2004). Genest and McKay (1986) considered Archimedean copula C generated by ϕ , then

$$\tau_{Arch} = 4 \int_0^1 \frac{\phi(t)}{\phi'(t)} dt + 1,$$

where τ_{Arch} is Kendall's tau of Archimedean copula C . Kendall's tau and Spearman's rho of nine copulas are provided in Table 2. Note that $D_m(\theta) = (1/\theta^m) \int_0^\theta t^m / (e^t - 1) dt$ in Frank copula.

4. Application of copula to multivariate control chart

One of the primary multivariate attribute schemes for control chart was proposed by Patel (1973) which was based on Hotelling-type χ^2 statistic for monitoring multivariate binomial or multivariate Poisson distributions. Considering adequate sampling subgroup sizes, the data vectors to follow a multivariate normal distribution are assumed. Therefore, the concept of Shewhart based control charting for multivariate normal observations are used. Marcucci (1985) applied a multinomial distribution to develop a control chart for the case that proportions in each quality category were known or estimated using a base period. Lu (1998) used a χ^2 statistic (a weighted nonconformities counts) and applied it in a multivariate Shewhart np-chart (Dokouhaki and Noorossana, 2013).

There have been several recent papers published using copulas on the control chart. Larpiatataworn (2003) proposed an artificial neural network (ANN) for bivariate binomial monitoring. He used the assumptions of adequate sample size and existence of positive correlation. He discussed different values of correlation between two attribute characteristics.

A copula based bivariate zero inflated Poisson distribution to monitor rare events was proposed by Fatahi *et al.* (2011). They found that one difficulty in developing multivariate attribute control charts was the lack of the related joint distribution. So, then a bivariate (or multivariate) attribute control chart can be developed based on types I and II errors if it was possible to generate the joint

Table 2: Kendal’s tau and Spearman’s rho of nine different copula functions

Copula	Kendal’s tau	Spearman’s rho	θ – domain
Normal	$\frac{\arcsin(\theta)}{\pi/2}$	$\frac{\arcsin(\theta/2)}{\pi/6}$	$[-1, 1]$
Frank	$1 - \frac{4(D_1(\theta) - 1)}{\theta}$	$1 - \frac{12(4D_1(\theta) - 2D_2(\theta))}{\theta}$	$(-\infty, \infty)$
Gumbel	$\frac{\theta - 1}{\theta}$	-	$[1, \infty)$
Clayton	$\frac{\theta}{\theta + 2}$	-	$(0, \infty)$
AMH	$\frac{\arcsin(\theta)}{\pi/2}$	-	$[-1, 1]$
FGM	$\frac{2\theta}{3 - \theta}$	$\frac{\theta}{3}$	$[-1, 1]$
Joe	$\frac{2\theta}{9}$	$\frac{\theta}{3}$	$[1, \infty)$
Plackett	-	$\frac{\theta + 1}{\theta - 1} - \frac{2\theta \ln \theta}{(\theta - 1)^2}$	$[0, \infty)$
Raftery	$\frac{2\theta}{3 - \theta}$	$\frac{\theta(4 - 3\theta)}{(2 - \theta)^2}$	$[0, \infty)$

distribution of two (or more) attribute characteristics. The copulas function is a solution to the matter. In their article, applying the copula function approach achieved the joint distribution of two correlated zero inflated Poisson (ZIP) distributions.

Considering X_1 and X_2 as two ZIP random variables with parameters θ_1, μ_1 and θ_2, μ_2 , respectively, and the correlation coefficient ρ , they derived a copula function which best fits the related joint distribution function. To do so, from the probability distribution function of Poisson distribution, the marginal CDFs are calculated as:

$$F(X_1) = P(X_1 \leq x_1) = \begin{cases} 1 - \theta_1 + \theta_1 \times e^{-\mu_1}, & x_1 = 0, \\ 1 - \theta_1 + \theta_1 \sum_{t=0}^{x_1} \frac{\mu_1^t}{t!}, & x_1 = 1, 2, \dots, \end{cases}$$

$$F(X_2) = P(X_2 \leq x_2) = \begin{cases} 1 - \theta_2 + \theta_2 \times e^{-\mu_2}, & x_2 = 0, \\ 1 - \theta_2 + \theta_2 \sum_{t=0}^{x_2} \frac{\mu_2^t}{t!}, & x_2 = 1, 2, \dots \end{cases}$$

There marginal cdfs and the correlation coefficient ρ are substituted into the copula function (Shih and Louis, 1995), to define the related copula function which best fits the joint distribution. This joint distribution, is then used to develop a bivariate control chart that can be used to monitoring the correlated rare events. The copula-based bivariate ZIP control chart is compared with the simultaneous use of two separate univariate ZIP control charts. Based on the ARL measure, it is shown that the proposed control chart is better than the simultaneous use of two separate univariate charts. In addition, a real case study related to the environmental air in a sterilization process is investigated to show the applicability of the developed control chart. Monitoring multivariate (or bivariate) rare events, where correlations between the two attribute quality characteristics exist, is an important issue in statistical process control. There are few methods in the literature of monitoring such bivariate quality characteristics applying a multivariate (or bivariate) distribution of the attributes. According

to the developed copula, based on a bivariate distribution, they contributed a bivariate ZIP control chart. The ARL criterion indicates that this chart has better performance than the simultaneous use of two separate ZIP charts.

A copula based on the multivariate Shewhart control charts were designed under the assumption of independence of consecutive observations. The most popular multivariate control chart to monitor the mean of a distribution is probably the Hotelling T^2 rule. Unfortunately, this rule relies on the assumption that the distribution under control is Gaussian, which is rarely true in practice. In the presence of dependence the authors usually assume dependencies in the form of autocorrelated and normally distributed data. However, there exist many types of dependencies that are described by other mathematical models. The question arises then, on how robust classical control charts are for different types of dependencies. This problem has been sufficiently discussed for the concept of cases involving autocorrelated and normal data. The concept of copulas to model dependencies of other data types are used for efficient control charts.

Hryniewicz and Szewi (2010) have shown a new control chart based on non-parametric Kendall's tau statistics by a bivariate normal distribution. In 2012, Hryniewicz discussed the robustness of Shewhart control charts in the case of dependent data described by autocorrelated and normal data. He used the concept of copula to model dependences and the Monte Carlo simulation experiments to investigate the impact of type and strength of dependence in data on the value of the ARL of Shewhart control charts. His results confirmed the findings of many authors who considered the behavior of Shewhart control charts for cases of dependent data described by an autoregressive stochastic process. A new finding was that the demonstration that the type of dependence, encapsulated in the type of respective copula, plays important role. Moreover, it becomes clear that the knowledge of the strength of dependence, measured using popular statistical measures of dependence such as Kendall's tau is insufficient to evaluate the properties of the Shewhart control chart. From the results presented in their paper, one can derive the following recommendations. First, it is necessary to detect the existence of dependence in data. Then it is necessary to indicate the copula that fits to the observed data. Therefore, the needs to propose even approximate investigations on finding the appropriate corrections to classical procedures, which is similar to those proposed in the cases of dependencies described by the normal copula.

In 2013, Verdier presented a new approach for non-normal multivariate cases, including Cauchy, Student's t and normal distributions. He proposed new copula function for the Hotelling's T^2 control chart based on the data depth, by the construction of a tolerance region obtained from a density level set estimation. The method follows a plug-in approach in which the density of the observations is previously estimated. This estimation conducts inducing copulas modelling, that is increasingly popular to the multivariate modelling. The Hotelling T^2 control chart is probably the most used rule in industry for multivariate fault detection, which relies on the assumption that the observations under control are Gaussian. When the T^2 is applied on non-normal multivariate observations, it can lead to many false alarms and non-detections. Alternative approaches have been proposed based on parametric or non-parametric methods. The control chart obtained is not subject to the curse of dimension for non-parametric approaches and remains relatively flexible thanks to a wide choice of copula families (Verdier, 2013). However, they still must consider the problem of deciding if an observation X_{test} has been generated from a reference distribution F . If the objective is to detect a change in dynamic system (from an unknown time t_c , the system is out of control and all the new observation are generated from a new distribution G), the control chart which proposed can be called without memory like the Hotelling T^2 rule or the data depth rule. At the time step t , the decision of declaring the system in-control or out-of-control is taken only using the observation X_t (and obviously

the learning sample).

For this, they defined $\hat{F}_{h_n(X)}$ the distribution of a random sample of $h_n(X)$. This sample is noted $h_n(X_{1,n}), \dots, h_n(X_{N,n})$ and they have

$$\hat{F}_{h_n(X)}(y) = \frac{1}{N} \sum_{i=1}^N \{\mathbf{I}(h_n(\mathbf{X}_{i,n}) \leq \mathbf{y})\},$$

where $\mathbf{I}(\cdot)$ is the indicator function. Therefore, the estimate of C_α is obtained by

$$\hat{C}_{N,n} = \text{Inf} \{c \in \mathbf{R} : \hat{F}_{h_n(X)}(c) \geq \alpha\}.$$

The control charts without memory are particularly well suited for the detection of large changes in the distribution F . However, the control charts without memory, like the MEWMA (a combination of the T^2 and an EWMA procedure (Montgomery, 2013)) for Gaussian observations or nonparametric approaches (Qiu, 2008; Zou and Tsung, 2011) will be preferable to deal with a small change (for example a small variation of the process mean). They also commented for future work to combine their method with (for example) an EWMA procedure in order to take into account the history of the process (the observations X_{t-1}, X_{t-2}, \dots) and to obtain a more efficient control chart in the case of a small change.

Despite the details mentioned in previous papers, differences in distribution can be found in control charts. In a manufacturing process, time represents an attribute or variable measure observed as consecutive events of concern. The distribution of the time for an event is known as exponential distribution when the probability of an event in the next small time interval does not vary though time. It is a continuous distribution and widely used to monitor time successive occurrences of events. For studies on the MCUSUM control chart, Busaba *et al.* (2012) used an integral equation approach to derive explicit formulas for the ARL (the first passage times) for CUSUM when observations are negative exponential distributed. Simulations are conducted to compare the performance of the explicit formulas with numerical approximations. The computing time for the explicit formulas is found to be approximately 10 seconds, which is significantly less than the computational time required for numerical approximations. Explicit formulas of ARL, for the CUSUM charts when the observations have negative exponential distribution have been derived analytically. The accuracy of these explicit expressions has been checked by a comparison with numerical solutions of integral equations based on using Gauss-Legendre integration rules to approximate integrals. The numerical results and the values from the explicit formulas were in excellent agreement. The computation times required for the numerical computations were approximately 40–50 minutes compared with less than 1 second for the explicit formulas as show in Table 3.

Petcharat *et al.* (2014) applied the Fredholm type integral equations method to derive explicit formula of the ARL for a CUSUM chart, when observations are described by a first order moving average MA(1) process, with exponential white noise. They compared the computational time between our analytical explicit expressions for the ARL performance with the one obtained via the Gauss-Legendre numerical scheme for integral equations. They found that those methods are in excellent agreement; however, the computational time of the former takes approximately 1 second while the latter method consumes an approximate computational time of 11 minutes. They presented the explicit formula to evaluate the ARL_0 and ARL_1 of the CUSUM chart when observations are the first order moving average process, MA(1) with exponential white noise distribution. The accuracy and computational speed for analytical formula was compared with values obtained from the numerical integration method by Fredholm integral equation (Rao *et al.*, 2001). They have shown that analytical

Table 3: Comparisons of values from explicit formulas $j(x)$ and numerical approximations $j^*(x)$ for ARL and Time used for fixed $b = 0.69$ and $a = 0$ with $m = 800$ for CUSUM chart

λ	$x = 0$			$x = 0.5$		
	$j(x)$	$j^*(x)$	Diff(%)	$j(x)$	$j^*(x)$	Diff(%)
1.00	207.4500	207.4700 3,135.34 ^a	0.0096	207.8440	207.8640 2,683.63	0.0096
1.01	65.3108	65.3171 2,819.28	0.0096	65.7013	65.7077 4,826.59	0.0097
1.05	17.9863	17.9879 2,818.59	0.0089	18.3651	18.3669 2,842.44	0.0098
1.07	13.3735	13.3747 2,805.44	0.0090	13.7468	13.7481 2,843.09	0.0095
1.10	9.6620	9.7670 2,818.70	0.0082	10.1315	18.3669 2,823.83	0.0089
1.20	5.3939	5.3942 2,838.19	0.0056	5.7346	5.7351 2,825.34	0.0087
2.00	1.8146	1.8146 2,838.69	0.0000	2.0358	2.0359 2,850.14	0.0049
3.00	1.3874	1.3874 2,886.66	0.0000	1.5409	1.5410 2,871.03	0.0065
4.00	1.2514	1.2514 2,837.50	0.0000	1.3689	1.3690 2,873.27	0.0073

^a = CPU's time used (second).

expressions are very accurate in the same manners of the numerical integration method; however, the latter method is very time consuming in terms of computational time.

Dokouhaki and Noorossana (2013) presented a copula Markov CUSUM chart, based on the Markov CUSUM chart that is good at monitoring the auto-correlation statistical process control and for monitoring two correlated binomial data sets with an auto-correlation. There is numerous research related to the multivariate variables control charting (focusing on the variable quality characteristics) in the statistical process control (SPC) literature; however, fewer investigations could be found regarding the multivariate attributes control charts (relying on the attribute quality characteristics). Multivariate attribute control charting makes it interesting to monitor the auto-correlated data, since the real-world processes usually include the data based on an auto-correlation structure. Ignoring the auto-correlation structure in developing a multivariate control chart increases the type I and type II errors simultaneously and consequently reduces the performance of the chart. The most important difficulty in developing a multivariate attributes control charts is the absence of the joint distribution for quality characteristics. This deficiency can be dispelled through the use of a copula approach for developing joint distribution. In that paper, Dokouhaki and Noorossana use the Markov approach for modelling the auto-correlated data. Then, the copula approach is used to make the joint distribution of two auto-correlated binary data series. Finally, based on this joint distribution, They develop a CUSUM chart. Hence, the proposed chart is entitled the copula Markov CUSUM chart.

Suppose X and Y as two correlated binary data with an auto-correlated structure for each. Thus, the two correlated sequences of X_1, X_2, X_t, \dots and $Y_1, Y_2, \dots, Y_t, \dots$, each with the value of 0 or 1, denote the conforming or non-conforming observations of two characteristics of a process. The auto-correlation coefficients related to X and Y and the correlation are show by ρ, ρ' , and $\rho_{x,y}$, respectively.

Consider p as the probability that an item is defective and p which is achieved by the process through the lack of any assignable causes of the process variation. Hence, p_0 and p'_0 are the in-control values related to X and Y , respectively which detect any changes in the study. Also, p_1 and p'_1 are the

out-of-control detectable values that they expect in the control chart alarm.

The transition probability of the two-state Markov model regarding the variable X (or Y) is a conditional probability and would be as:

$$p_{ij} = P(X_t = j | X_{t-1} = i), \quad t = 1, 2, \dots; \quad p'_{ij} = P(Y_t = j | Y_{t-1} = i), \quad t = 1, 2, \dots$$

Hence, for a typical variable like X whose value are 0 or 1, four transition probabilities can be derived as:

$$p_{00} = 1 - p(1 - \rho), \quad p_{01} = p(1 - \rho), \quad p_{10} = (1 - p)(1 - \rho), \quad p_{11} = 1 - (1 - p)(1 - \rho).$$

In addition, the estimation of p (as the proportion) and ρ (as the correlation between two successive observations) are

$$\hat{p} = P(X_t = 1) = \frac{p_{01}}{p_{01} + p_{10}} \quad \hat{\rho} = 1 - (p_{01} + p_{10}).$$

They used the above equations to develop a univariate control chart entitled Markov-based CUSUM which shows a better performance than other related univariate charts. However, the conditional joint distribution of the variables X and Y needs to be known in order to use the Markov approach in a bivariate control chart. They use the copula function to develop such a distribution since the conditional joint distribution of the two variables cannot be achieved straightforwardly.

In order to derive a bivariate copula function, denote the two cdfs $F(x_1)$ and $F(x_2)$ by the two uniform random variables U and V , respectively. If and are independents, they have copula function $C(u, v) = uv$. If the two random variables are perfectly correlated, we have the copula function $C(u, v) = \min(uv)$. Mixing these two copula functions by a correlation coefficient (ρ_{x_1, x_2}) , we have new copula function as follows (Shih and Louis, 1995):

$$C(u, v) = \begin{cases} (1 - \rho_{x_1, x_2})uv + \rho_{x_1, x_2} \min(u, v), & \rho_{x_1, x_2} > 0, \\ (1 + \rho_{x_1, x_2})uv + \rho_{x_1, x_2}(u - 1 + v)\Theta(u - 1 + v), & \rho_{x_1, x_2} \leq 0, \end{cases}$$

where $\Theta(a) = \begin{cases} 1, & a \geq 0, \\ 0, & a < 0. \end{cases}$ There are considerable number of multivariate variables control charts (focusing on the variable quality characteristics), but fewer control charts could be found regarding the multivariate attribute characteristics (relying on attribute quality characteristics). Usually, in control charting, it is considered that data are independent, but the real-world processes usually include data based on an auto-correlation structure. Ignoring the auto-correlation structure in developing a multivariate control chart increases the two statistical type errors simultaneously and reduces the performance of the chart.

Based on the presented concept above, in order to create a conditional copula function, first you should generate the marginal cumulative functions. Since there are auto-correlation structures for X and Y , you should consider U and V equal to $F(x_t|x_{t-1})$ and $F(y_t|y_{t-1})$, respectively, as the following:

$$F(x_t|x_{t-1}) = \begin{cases} 0, & x_t < 0, \\ p_{00}^{(1-x_{t-1})} p_{10}^{x_{t-1}}, & 0 \leq x_t < 1, \\ p_{00}^{(1-x_{t-1})} p_{10}^{x_{t-1}} + p_{01}^{(1-x_{t-1})} p_{11}^{x_{t-1}} = 1, & x_t \geq 1, \end{cases}$$

$$F(y_t|y_{t-1}) = \begin{cases} 0, & y_t < 0, \\ p_{00}^{(1-y_{t-1})} p_{10}^{y_{t-1}}, & 0 \leq y_t < 1, \\ p_{00}^{(1-y_{t-1})} p_{10}^{y_{t-1}} + p_{01}^{(1-y_{t-1})} p_{11}^{y_{t-1}} = 1, & y_t \geq 1. \end{cases}$$

For $\rho_{x,y} > 0$ and $\rho_{x,y} \leq 0$, the copula function can be computed from the following equations:

$$\underbrace{C_{u,v}}_{\text{for } \rho_{x,y} > 0} = \begin{cases} 0, & x_t < 0, y_t < 0, \\ (1 - \rho_{x,y}) \left(p_{00}^{(1-x_{t-1})} p_{10}^{x_{t-1}} \right) \left(p'_{00}^{(1-y_{t-1})} p'_{10}^{y_{t-1}} \right) \\ \quad + \rho_{x,y} \min \left\{ \left(p_{00}^{(1-x_{t-1})} \right), \left(p'_{00}^{(1-y_{t-1})} p'_{10}^{y_{t-1}} \right) \right\}, & 0 \leq x_t < 1, 0 \leq y_t < 1, \\ (1 - \rho_{x,y}) \left(p_{00}^{(1-x_{t-1})} p_{10}^{x_{t-1}} \right) (1) + \rho_{x,y} \min \left\{ \left(p_{00}^{(1-x_{t-1})} p_{10}^{x_{t-1}} \right), 1 \right\}, & 0 \leq x_t < 1, y_t \geq 1 \\ (1 - \rho_{x,y}) (1) \left(p'_{00}^{(1-y_{t-1})} p'_{10}^{y_{t-1}} \right) + \rho_{x,y} \min \left\{ 1, \left(p'_{00}^{(1-y_{t-1})} p'_{10}^{y_{t-1}} \right) \right\}, & x_{t-1} \geq 1, 0 \leq y_t < 1, \\ 1, & x_t \geq 1, y_t \geq 1, \end{cases}$$

$$\underbrace{C_{u,v}}_{\text{for } \rho_{x,y} \leq 0} = \begin{cases} 0, & x_t < 0, y_t < 0, \\ (1 + \rho_{x,y}) \left(p_{00}^{(1-x_{t-1})} p_{10}^{x_{t-1}} \right) \left(p'_{00}^{(1-y_{t-1})} p'_{10}^{y_{t-1}} \right) + \rho_{x,y} \left(p_{00}^{(1-x_{t-1})} \right) - 1 \\ \quad + \left(p'_{00}^{(1-y_{t-1})} p'_{10}^{y_{t-1}} \right) \times \Theta \left(p_{00}^{(1-x_{t-1})} p_{10}^{x_{t-1}} - 1 + p'_{00}^{(1-y_{t-1})} p'_{10}^{y_{t-1}} \right), & 0 \leq x_t < 1, 0 \leq y_t < 1, \\ (1 + \rho_{x,y}) \left(p_{00}^{(1-x_{t-1})} p_{10}^{x_{t-1}} \right) + \rho_{x,y} \left(p_{00}^{(1-x_{t-1})} p_{10}^{x_{t-1}} \right) \times \Theta \left(p_{00}^{(1-x_{t-1})} p_{10}^{x_{t-1}} \right), & 0 \leq x_t < 1, y_t \geq 1, \\ (1 + \rho_{x,y}) \left(p'_{00}^{(1-y_{t-1})} p'_{10}^{y_{t-1}} \right) + \rho_{x,y} \left(p'_{00}^{(1-y_{t-1})} p'_{10}^{y_{t-1}} \right) \times \Theta \left(p'_{00}^{(1-y_{t-1})} p'_{10}^{y_{t-1}} \right), & x_{t-1} \geq 1, 0 \leq y_t < 1, \\ 1, & x_t \geq 1, y_t \geq 1, \end{cases}$$

where $\Theta(a) = \begin{cases} 1, & a \geq 0, \\ 0, & a < 0. \end{cases}$

In order to develop a multivariate chart regarding the auto-correlation, they use the Markov approach to the auto-correlated data. The copula approach is then used to make the joint distribution of the two auto-correlated binary data series.

Kuvattana *et al.* (2015b) proposed four types of copulas on the CUSUM control chart when observations have exponential distribution, which the MCUSUM control chart statistics are

$$Y_t = \left[S'_t \Sigma^{-1} S_t \right]^{\frac{1}{2}}, \quad t = 1, 2, 3, \dots$$

The signal gives an out-of-control if $Y_t > h$ where h is the control limit, and use $H(x, y) = C(F(x), F(y); \theta)$ with a copula $C : [0, 1]^2 \rightarrow [0, 1]$ where θ is a copula parameter. The level of dependence is measured by Kendall's tau ($-1 \leq \tau \leq 1$). For moderate and strong dependence, Kendall's tau values are defined as 0.5 and 0.8, respectively. They used the Monte Carlo simulation to investigate the ARL value. The dependence of random variables are used and measured by Kendall's tau in each copula. The numerical results show that negative dependence normal copula is superior. For positive dependence, in the case of two parameter shifts, normal copula is superior and Gumbel copula is superior in the case of one parameters shift. Dependence measures of two or more variables can be investigated in terms of various copulas. Their paper show CUSUM control chart for four types of copulas and level of dependence are measured by Kendall's tau values. Table 4 shows that normal copula is superior for negative dependence as well as that normal copula is superior for positive dependence in the lower case of the shift in two of parameters; however, Gumbel copula is superior in the case of the shift in one of parameters.

For the study on MEWMA control chart. Xie *et al.* (2011) presented exponential distribution commonly encountered in the manufacturing process, reliability analysis, and human-service management. Time between events (TBE) charts have been suggested to monitor exponential data. However,

Table 4: Efficiency of copula depends on positive and negative dependences

Parameters	Positive dependence		Negative dependence	
	Moderate	Strong	Moderate	Strong
One parameter	Normal	Normal	Normal	Clayton
Two parameters	Gumbel	Gumbel	Normal	Normal

existing studies on TBE charts are limited to univariate cases that assume one process characteristic of interest. In this paper, two MEWMA charts (one based on the raw observations and the other based on the transformed data) are proposed for the simultaneous monitoring of the mean vector of Gumbel’s bivariate exponential (GBE) and TBE model. A numerical example is provided to illustrate the implementation of the two MEWMA charts. They compare the ARL performances of the two proposed charts with the following individual TBE charts pairs: paired individual t charts, paired individual EWMA charts on raw data, and paired individual EWMA charts on transformed data. The results of the comparative studies show that our MEWMA charts outperform all other charts. The proposed MEWMA charts can be easily extended to higher dimensions. Some brief discussions concerning the monitoring of Gumbel’s multivariate exponential model with more than two variables are also.

In Kuvattana *et al.* (2016) proposed four types of copulas on the EWMA control chart when observations are from an exponential distribution using a Monte Carlo simulation approach. The performance of the control chart is based on the ARL which is compared for each copula. Copula functions for specifying dependence between random variables are used and measured by Kendall’s tau. The results revealed that it is necessary to detect the dependence of the observation to indicate the copula that fits the observation. For moderate dependence, the normal copula can be used for almost all shifts.

5. Performance characteristics to multivariate control chart

The most frequently characteristic that describes the performance of control charts is the ARL, which is the average number of points that must be plotted before a point indicates an out-of-control condition. The ARL is classified into ARL_0 and ARL_1 . Theoretically, an acceptable ARL in a control condition should be adequate and is called ARL_0 . However, ARL should be small when the process is out-of-control, in which case it is called the average of delay time (ARL_1). The process is in-control condition when $ARL_0 = 1/\alpha$ and where α is the probability that Hotelling’s T^2 statistic exceeds UCL and $ARL_1 = 1/(1 - \beta)$ where β is the probability of the out-of-control condition (Bersimis *et al.*, 2007).

Figure 1 shows the ARL of the Hotelling’s T^2 control chart with Kendall’s tau values equal 0.8 when the shifts are in one of the exponential parameters (Sukparungsee *et al.*, 2016). Parameters (μ_1, μ_1) representing (1, 1.5), (1, 2), (1, 2.5) and (1, 3) are given as the sequence order in x -label, respectively. It shows that the normal copula is adequate for all shifts.

The ARL as the expected number of sampling subgroups until an out-of-control signal occurs is the most commonly used performance measure. However, the ARL would not be suitable measure when control charts with different sampling subgroup sizes are being compared. In such situations, the ANOS is a positive measure to be used, see more Niaki and Nasaji (2011); Dokouhaki and Noorossana (2013). They used ANOS rather than ARL, since their chart is based on individual observations and chart work (according to subgroups) greater than one observation.

In this review paper, the result of the Monte Carlo simulation in R statistical software with the number of simulation runs 50,000 and sample size 10,000 is presented. Observations were from an

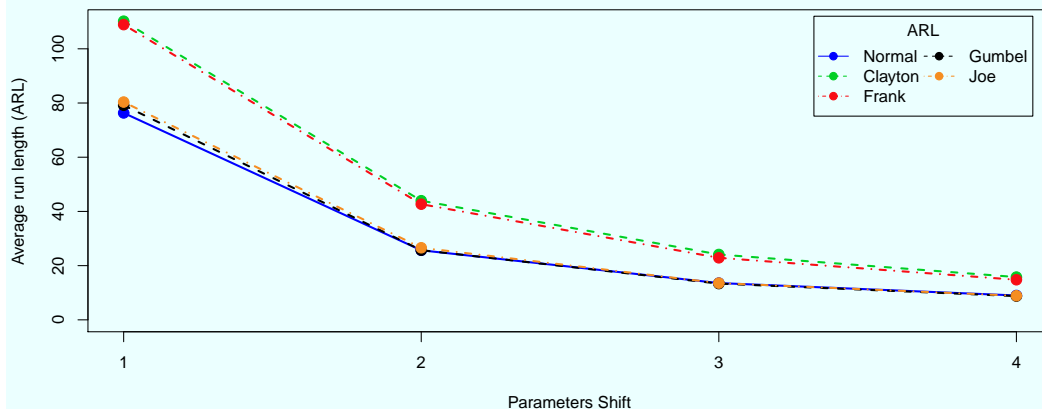


Figure 1: Average run length of the Hotelling’s T^2 control chart with Kendall’s tau values equal 0.8 when the shifts are in one of the exponential parameters.

Table 5: Average run length of control charts with Kendall’s tau equal to 0.8 in the case of one parameter shift

Parameters		ARL ₀ and ARL ₁					
μ_1	μ_2	Multivariate Shewart		MCUSUM		MEWMA	
		Normal	Frank	Normal	Frank	Normal	Frank
1	1.00	369.873	370.159	369.860	370.129	370.053	369.053
1	1.04	336.555	340.471	336.503	339.594	337.219	339.662
1	1.06	320.981	328.888	321.741	325.292	322.442	326.628
1	1.08	304.142	315.771	304.586	308.313	305.674	312.215
1	1.50	76.343	108.924	85.849	105.377	68.434	91.958
1	2.00	25.663	42.676	25.383	36.230	15.645	23.565
1	2.50	13.568	22.887	11.377	16.566	5.504	8.758
1	3.00	8.986	14.788	6.387	9.336	2.423	3.870

exponential distribution with in-control parameter $\alpha = 1$. The shift size is reported in terms of the quantity $\delta = \mu - \mu_0$ and large value of δ correspond to bigger shift in the mean. The value $\delta = 0$ and the process mean $\mu = 1$ are in-control. The process means are 1.04, 1.06, 1.08, 1.5, 2, 2.5 and 3 for out-of-control. The simulation result were conducted to evaluate the performance of the multivariate Shewhart control chart, the MCUSUM control chart and the MEWMA control chart with $\lambda = 0.05$. Copulas are restricted to cases of positive and negative dependence. For all copula model, the setting θ corresponds with Kendall’s tau. The level of dependence is measured by Kendall’s tau values ($-1 \leq \tau \leq 1$) which are defined to 0.8 and -0.8 , respectively (Kuvattana *et al.*, 2015a).

The results are presented in Tables 5–7, and the different values of exponential parameters denote μ_1 for the variable X and μ_2 for variable Y . For in-control, the three control charts were chosen by setting the desired $ARL_0 = 370$ for each copula. Tables 5 and 6 show strong positive dependence ($\tau = 0.8$), while Table 7 and Figure 2 show strong negative dependence ($\tau = -0.8$).

Table 5 shows the mean shift of $\tau = 0.8$ when μ_1 is fixes at 1. For small shift ($\mu_1 = 1, 1.02 \leq \mu_2 \leq 1.06$), the ARL_1 values of the normal copula on the MCUSUM control chart are less than the other copulas. For moderate shift ($\mu_1 = 1, 1.08 \leq \mu_2 \leq 1.5$), the ARL_1 value of the normal copula on the multivariate Shewhart control chart are less than other copulas. And for large shifts ($\mu_1 = 1, 2 \leq \mu_2 \leq 3$), the ARL_1 value of the normal copula on the multivariate MEWMA control chart are less than other copulas.

Table 6: Average run length of control charts with Kendall’s tau equal to 0.8 in the case of two parameter shift

Parameters		ARL ₀ and ARL ₁					
		Multivariate Shewart		MCUSUM		MEWMA	
μ_1	μ_2	Normal	Frank	Normal	Frank	Normal	Frank
1.00	1.00	369.873	370.159	369.860	370.129	370.053	369.885
1.02	1.02	340.421	344.763	342.224	338.785	341.084	343.795
1.06	1.06	288.406	298.431	292.618	290.538	291.095	293.816
1.08	1.08	266.017	280.742	304.586	308.313	269.346	271.621
1.50	1.50	68.052	81.951	76.813	80.001	59.246	67.359
2.00	2.00	27.244	34.117	27.655	31.957	16.999	21.105
2.50	2.50	15.157	19.408	13.508	16.533	6.765	9.169
3.00	3.00	10.014	13.102	7.911	10.005	3.145	4.846

Table 7: Average run length of control charts with Kendall’s tau equal to -0.8 in the case of two parameter shift

Parameters		ARL ₀ and ARL ₁					
		Multivariate Shewart		MCUSUM		MEWMA	
μ_1	μ_2	Normal	Frank	Normal	Frank	Normal	Frank
1.00	1.00	370.012	369.959	369.840	369.898	370.103	370.002
1.04	1.04	316.280	317.664	319.293	318.649	319.514	320.244
1.06	1.06	293.561	293.140	296.670	296.638	297.902	297.923
1.08	1.08	272.739	268.715	276.556	276.237	277.414	277.819
1.50	1.50	74.906	74.165	80.594	81.915	65.401	67.597
2.00	2.00	30.636	30.312	28.865	29.362	18.598	19.034
2.50	2.50	17.430	17.194	14.262	14.522	7.256	7.421
3.00	3.00	11.644	11.515	8.110	8.190	3.435	3.509

In Table 6, in the cases of the same shift in both exponential parameters of $\tau = 0.8$, the ARL₁ values of the Frank copulas and of the normal copulas on the MCUSUM control chart are less than the other copulas for a small shift. The ARL₁ value of the normal copula on the multivariate Shewart control chart are less than the other copulas for moderate shift. And the ARL₁ values of the normal copula on the MEWMA control chart are less than the other copulas for large shift.

Figure 2 shows the ARL of MCUCUM control chart with Kendall’s tau equal to -0.8 in the case of one parameter shift. Parameters (μ_1, μ_1) representing (1, 1), (1, 1.02), (1, 1.06), (1, 1.08), (1, 1.5) and (1, 2) are given as the sequence order in x-label, respectively. In the cases of the same shift in both exponential parameters of $\tau = -0.8$, the ARL₁ values of the Frank copulas and of the normal copulas on the MCUSUM control chart are less than other copulas for a small shift, the ARL₁ values of the normal copula on the multivariate Shewart control chart are less than other copulas for a moderate shift, and ARL₁ values of the normal copula on the MEWMA control chart are less than the other copulas for a large shift.

6. Discussion and conclusion

The discussion about the dependence measurements of two or more variables can be investigated in terms of various copulas modelling for MSPC , especially in multivariate control charts. Some copulas can be useful to analyze the dependence structure of data. Another problem with the application of the procedures designed to control autocorrelated data is the knowledge of the structure of correlation. Many practitioners have commented that MSPC with dependent data is very difficult, especially for not well-trained in statistics practitioners who need efficient tools to discriminate between complicated problems with dependent data and relatively simple problems when the observed data are independent.

In this review paper, the results of three control charts like multivariate Shewart, MCUSUM and

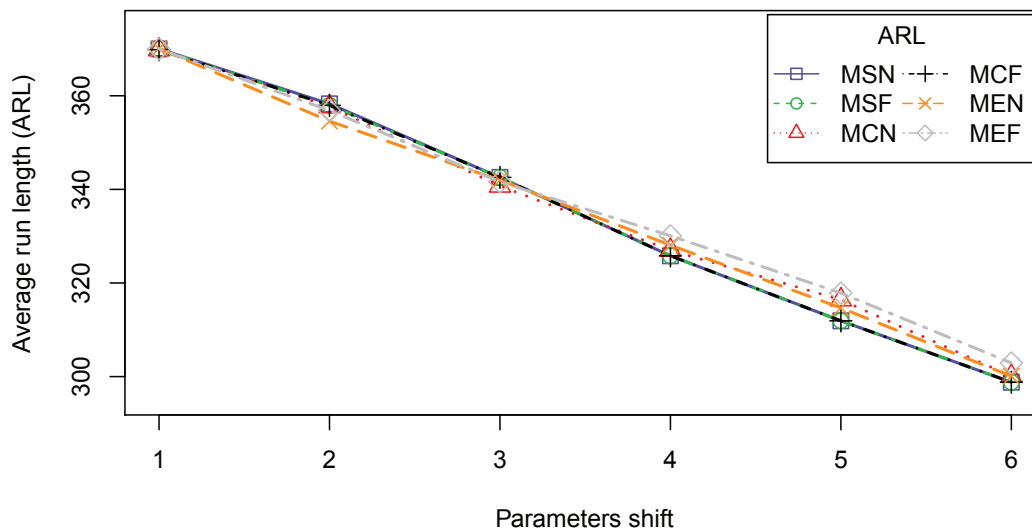


Figure 2: Average run length of MCUSUM control chart with Kendall's tau equal to -0.8 in the case of one parameter shift.

MEWMA, and the use of copulas modelling for a bivariate case are considered. The performance of control charts, based on the ARL are compared. The results revealed that the MCUSUM and the MEWMA control charts perform better than the multivariate Shewhart control chart for small shifts while the multivariate Shewhart is superior to the MCUSUM and the MEWMA control charts for moderate shifts, and the MEWMA control chart is more effective for large shifts. Families of well-known copulas are considered and compared for the multivariate control charts and process control, in this review. The results show that the MCUSUM control chart for three families of copulas are useful and the recommended level of dependence is measured by Kendall's tau values. The results revealed the necessity to detect the observation dependence to indicate the copula and use that detection to fit the observation of the MCUSUM chart for small and large dependencies.

The varieties of copula including vine copula and extreme copula can be designed on the same line for MSPC on further research. We can apply more copulas including asymmetric ones to simultaneously measure more than two variables for MSPC. Furthermore, the construction of copula for a multivariate control chart are required to dealing with correlated data, and future investigations should concentrate on finding the appropriate correction to classical procedures that are similar to those proposed in the case of dependencies as described by the normal copula.

Acknowledgements

This work was supported by the Mahasarakham University, Thailand.

References

- Alves CC, Samohyl RW, and Henning E (2010). Application of multivariate cumulative sum control charts (MCUSUM) for monitoring a machining process. In *Proceedings of the 16th International Conference on Industrial Engineering and Operations Management*, Sao Carlos, Brazil.

- Bersimis S, Panaretos J, and Psarakis S (2005). Multivariate statistical process control charts and the problem of interpretation: a short overview and some applications in industry. In *Proceedings of the 7th Hellenic European Conference on Computer Mathematics and its Applications*, Athens, Greece.
- Bersimis S, Psarakis S, and Panaretos J (2007). Multivariate statistical process control: an overview, *Quality and Reliability Engineering International*, **23**, 517–543.
- Bourke PD (1991). Detecting shift in fraction nonconforming using run-length control charts with inspection, *Journal of Quality Technology*, **23**, 225–238.
- Busaba J, Sukparungsee S, Areepong Y, and Mititelu G (2012). Analysis of average run length for CUSUM procedure with negative exponential data, *Chiang Mai Journal of Science*, **39**, 200–208.
- Crosier RB (1988). Multivariate generalizations of cumulative sum quality-control schemes, *Technometrics*, **30**, 291–303.
- Dokouhaki P and Noorossana R (2013). A copula Markov CUSUM chart for monitoring the bivariate auto-correlated binary observation, *Quality and Reliability Engineering International*, **29**, 911–919.
- El-Midany TT, El-Baz MA, and Abd-Elwahed MS (2010). A proposed framework for control charts pattern recognition in multivariate processes using artificial neural networks, *Expert Systems with Applications*, **37**, 1035–1042.
- Fatahi AA, Dokouhaki P, and Moghaddam BF (2011). A bivariate control chart based on copula function. In *Proceedings of the International Conference on Quality and Reliability (ICQR)*, Bangkok, Thailand, 292–296.
- Fatahi AA, Noorossana R, Dokouhaki P, and Moghaddam BF (2012). Copula-based bivariate ZIP control chart for monitoring rare events, *Communications in Statistics - Theory and Methods*, **41**, 2699–2716.
- Genest C and MacKay RJ (1986). The joy of copulas: bivariate distributions with uniform marginals, *The American Statistician*, **40**, 280–283.
- Hryniewicz O (2012). On the robustness of the Shewhart control chart to different types of dependencies in data, *Frontiers in Statistical Quality Control*, **10**, 19–33.
- Hryniewicz O and Szediw A (2010). Sequential signals on a control chart based on nonparametric statistical tests, *Frontiers in Statistical Quality Control*, **9**, 99–117.
- Joe H (1997). *Multivariate Models and Dependence Concepts*, Chapman & Hall, London.
- Joe H (2015). *Dependence Modeling with Copulas*, CRC Press, Boca Raton, FL.
- Khoo BC, Atta MA, and Phua HN (2009). A study on the performances of MEWMA and MCUSUM charts for skewed distributions. In *Proceedings of the 10th Islamic Countries Conference on Statistical Science*, Cairo, Egypt, 817–822.
- Kuvattana S, Sukparungsee S, Areepong Y, and Busababodhin P (2015a). Multivariate control charts for copulas modeling. In S. Ao, A. H. Chan, H. Katagiri, and L. Xu (Eds), *IAENG Transactions on Engineering Sciences: Special Issue for the International Association of Engineers Conferences 2015* (pp. 371–381), World Scientific Publishing, Singapore.
- Kuvattana S, Sukparungsee S, Busababodhin P, and Areepong Y (2015b). Efficiency of bivariate copulas on the CUSUM chart. In *Proceedings of the International Multiconference of Engineers and Computer Scientists (IMECS)*, Hong Kong.
- Kuvattana S, Sukparungsee S, Areepong Y, and Busababodhin P (2016). Bivariate copulas on the exponentially weighted moving average control chart, *Songklanakarin Journal of Science and Technology Preprint*, **38**, 569–574.
- Larpkiatataworn S (2003). *A neural network approach for multi-attribute process control with com-*

- parison of two current techniques and guidelines for practical use* (Ph.D. Thesis), University of Pittsburgh, PA.
- Lowry CA and Montgomery DC (1995). A review of multivariate control charts, *IIE Transactions*, **27**, 800–810.
- Lowry CA, Woodall WH, Champ CW, and Rigdon SE (1992). A multivariate exponentially weighted moving average control chart, *Technometrics*, **34**, 46–53.
- Lu XS (1998). Control chart for multivariate attribute processes, *Journal of Production Research*, **36**, 3477–3489.
- Marcucci M (1985). Monitoring multinomial processes, *Journal of Quality Technology*, **17**, 86–91.
- Mohmoud MA and Maravelakis PE (2013). The performance of multivariate CUSUM control charts with estimated parameters, *Journal of Statistical Computation and Simulation*, **83**, 721–738.
- Montgomery DC (2013). *Statistical Quality Control: A Modern Introduction* (7th ed), John Wiley & Sons, Singapore.
- Nayland College (2004). American new cars and truck, New Zealand.
- Nelsen RB (2016). *An Introduction to Copulas* (2nd ed), Springer, New York.
- Niaki STA and Nasaji SA (2011). A hybrid method of artificial neural networks and simulated annealing in monitoring auto-correlated multi-attribute processes, *International Journal of Advanced Manufacturing Technology*, **56**, 777–788.
- Patel HI (1973). Quality control methods for multivariate binomial and Poisson distributions, *Technometrics*, **15**, 103–112.
- Petcharat K, Areepong Y, Sukparungsee S, and Mititelu G. (2014). Exact solution for average run length of CUSUM Charts for MA(1) process, *Chiang Mai Journal of Science*, **41**, 1449–1456.
- Qiu P (2008). Distribution-free multivariate process control based on log-linear modelling, *IIE Transactions*, **40**, 664–677.
- Rao BV, Disney RL, and Pignatiello JJ (2001). Uniqueness and convergence of solutions to average run length integral equations for cumulative sum and other control charts, *IIE Transactions*, **33**, 463–469.
- Runger GC, Keats JB, Montgomery DC, and Scranton RD (1999). Improving the performance of the multivariate exponentially weighted moving average control chart, *Quality and reliability of Engineering International*, **15**, 161–166.
- Shih JH and Louis TA (1995). Inferences on the association parameter in copula models for bivariate survival data, *Biometrics*, **51**, 1384–1399.
- Sklar M (1959). Fonctions de répartition à n dimensions et leurs marges, *Publications de l'Institut de Statistique de l'Université de Paris*, **8**, 229–231.
- Sukparungsee S, Kuvattana S, Areepong Y, and Busababodhin P (2016). Bivariate copulas on the exponentially weighted moving average control chart, *Songklanakarin Journal of Science and Technology*, **38**, 569–574.
- Suriyakart W, Areepong Y, Sukparungsee S, Mititelu G (2012). Analytical Method of Average Run Length for Trend Exponential AR(1) Processes in EWMA Procedure, *IAENG International Journal of Applied Mathematics*, **42**, 250–253.
- Trivedi PK and Zimmer DM (2005). Copula modeling: an introduction for practitioners, *Foundations and Trends in Econometrics*, **1**, 1–111.
- Verdier G (2013). Application of copulas to multivariate control charts, *Journal of Statistical Planning and Inference*, **143**, 2151–2159.
- Xie M and Goh TN (1992). Some procedures for decision making in controlling high yield processes, *Quality and Reliability Engineering International*, **8**, 355–360.

Xie Y, Xie M, and Goh TN (2011). Two MEWMA charts for Gumbel's bivariate exponential distribution, *Journal of Quality Technology*, **43**, 50–56.

Zou C and Tsung F (2011). A multivariate sign EWMA control chart, *Technometrics*, **53**, 84–97.

Received November 16, 2016; Revised November 18, 2016; Accepted November 18, 2016