An assessment for the conditional performance of an SVDD-based chart

Anan Tang, Philippe Castagliola, Xuelong Hu, Fupeng Xie

An assessment for the conditional performance of an SVDD-based chart

Anan Tang*[1,2], Philippe Castagliola[3], Xuelong Hu[1,2], Fupeng Xie[4]

[1]Nanjing University of Posts and Telecommunications, Nanjing, China,
*Corresponding author
[2]Key Research Base of Philosophy and Social Sciences in Jiangsu–Information Industry Integration Innovation and Emergency Management Research Center, Nanjing, China
[3]Université de Nantes & LS2N UMR CNRS 6004, Nantes, France
[4]Nanjing University of Science and Technology, Nanjing, China

Abstract

The usual practice in Statistical Process Monitoring techniques assumes that the data distribution is known and the related parameters are accurately estimated. In practice, the underlying distribution and its parameters are rarely known, and control charts need to be constructed with parameters being estimated. Such issues have recently received an increasing attention in evaluating the properties of both parametric and nonparametric charts. However, the same study is seldom conducted for the control charts based on the data-driven tools. In this paper, we investigated the in-control performance of a nonparametric control chart based on the Support Vector Data Description (SVDD) theory. More specifically, we discuss the conditional effect of the training Phase-I samples on the Phase-II efficiency when different distributions are considered. Simulation results show that the conditional performance of the SVDD-based chart can be strongly affected by the Phase-I samples. It this situation, adjusted control limits with a specific number of available training sample is suggested.

Keywords: Support Vector Data Description; Control Chart; Phase-I; Phase-II; Conditional Performance

1 Introduction

Control charts as Statistical Process Monitoring (SPM) schemes play a crucial role in a variety of manufacturing and service fields, like, for example, in monitoring the non-conforming items in the production industry, the diseases outbreak in public-health surveillance, or the climate changes in meteorology. Some of these areas have drawn significant attention of many researchers, see Woodall (2006), Chen et al. (2015), Tanveer et al. (2019). In all of these situations, at least two essential characteristics are primarily recorded and the use of univariate control charts to monitor these quality characteristics separately is known to be misleading. Therefore, various multivariate schemes have been proposed for the simultaneous monitoring of correlated variables. The most frequently used multivariate charts, such as the Hotelling T^2 chart, the Multivariate Cumulative Sum (MCUSUM)

chart and the Multivariate Exponentially Weighted Moving Average (MEWMA) chart, are generally confined to the monitoring of the mean vector of multivariate normal processes.

However, in many applications, when the multivariate normality assumption is often questionable or the actual distribution is unknown, these parametric schemes may potentially be (highly) affected. For univariate processes, several studies have shown that a departure from normality severely deteriorates the monitoring properties of schemes based on the normal theory, see Graham et al. (2017), Asghari et al. (2018), Castagliola et al. (2018), Tang et al. (2019), Alevizakos et al. (2020). This situation is exacerbated for multivariate processes as the multivariate normality is even more uncommon than the univariate one. Therefore, in recent years, a host of nonparametric multivariate SPM schemes have emerged as attractive alternatives to multivariate parametric ones. For example, Zou and Tsung (2011) developed a multivariate EWMA scheme for monitoring process location by using a sign test as the statistic. Zou et al. (2012) further considered using the spatial ranks statistic, and they proposed a multivariate self-starting rank-based EWMA chart for monitoring process location. Holland and Hawkins (2014) developed a nonparametric scheme using an approximately distribution-free multivariate Wilcoxon-Mann-Whitney test. Mahmoud and Maravelakis (2013) and Dovoedo and Chakraborti (2016) investigated the effect of parameter estimation on the conditional performance of the multivariate CUSUM and EWMA charts, respectively. Huwang et al. (2019) studied the properties of the nonparametric EWMA chart for monitoring the shape matrix of a multivariate process. Readers are recommended to refer to Qiu (2018), Chakraborti and Graham (2019) for an extensive review of the literature on nonparametric charts. However, it is essential to note that, a significant part of the nonparametric schemes mention above still have some crucial prerequisites in describing the data, see Ning and Tsung (2013), Singh and Prasher (2019). For example, the continuity for the process variables is a common prerequisite in most of the nonparametric schemes focusing on the signs or the signed-ranks of the observations. Unfortunately, many processes do not meet such a prerequisite, and it is therefore highly desirable to propose SPM schemes for monitoring mixed-type data process (with both categorical and continuous variables) with different marginal distributions, see Ning and Tsung (2013).

As a boundary-based scheme, the Support Vector Data Description (SVDD) method only checks the distance between the new observations and the optimal boundary generated by the so called support vectors. Thus, SVDD-based control charts can generate an adaptable in-control region regardless of how the target data are distributed. Such flexibility makes it a highly active topic in modern nonparametric developments. Sun and Tsung (2003) first created a kernel distance-based chart (denoted as the K chart). Then, Kumar et al. (2006) further explored the over-fitting issues of the K chart in the presence of outliers. Sukchotrat et al. (2009) developed a robust parameter design strategy for a SVDD-based chart, whose control limits are established depending on the estimated percentile of the kernel distance. Gani et al. (2011) evaluated the performance of the K chart in a real industrial application, and the results showed that this K chart is more sensitive than the T^2 chart for detecting small shifts in the mean vector. Finally, Ning and Tsung (2013) investigated three design approaches for the parameter selection for a SVDD-based chart.

Both parametric and nonparametric SPM schemes are implemented in two stages: 1) in Phase-I, some preliminary analysis based on a training data set are performed, and in-control parameters are estimated, 2) in Phase-II, the in-control parameters obtained during the Phase-I are used to establish appropriate control limits for the monitoring of new incoming data. It is important to note that different Phase-I datasets (reference sample) will result in different parameter estimates. Consequently, the control limits and the chart's performance will vary across practitioners. An extensive amount

of researches has been done on parametric charts with estimated parameters, and most of which are about the monitoring of the mean, e.g. Saleh et al. (2015), Hu et al. (2018), the variance, e.g. Castagliola and Maravelakis (2011), Diko et al. (2017) and the median, e.g. Cheng et al. (2018), Tang et al. (2019). All of them have indicated that the use of estimated parameters seriously affects the in-control properties of the parametric control charts. Although many conditional designs of parametric control charts have been reported in other journals, relatively little attention has been devoted to the design of multivariate nonparametric control charts when the process parameters are unknown with few exceptions like Dovoedo and Chakraborti (2016). In the work of Dovoedo and Chakraborti (2016), they investigated the in-control performance of the MSEWMA sign chart in the unknown parameter case, and they found that an extremely large reference sample is needed to estimate parameters so that the chart can have identical performance as the known parameter case.

Clearly, the conditional performance topic has emerged as an important research area in the SPM literature. But, to the best of our knowledge, similar investigations have not been conducted for control charts based on the SVDD theory. The current research addresses this issue and it makes contributions in three aspects: (i) it investigates the impact of the chart's parameters on the detecting performance of the SVDD-based chart under various types of distributions, (ii) it evaluates how the number of Phase-I samples affects the Phase-II performance, and (iii) it proposes two percentile bootstrap-based methods to adjust the control limits.

The rest of this paper is organized as follows: In Section 2, the standard SVDD algorithm is presented. Section 3 introduces the corresponding SVDD-based control chart. The design of its parameters is also discussed, along with an illustrative example. In Section 4, we highlight the importance of considering the conditional performance when assessing the SVDD-based chart and two bootstrap methods are suggested to adjust the control limits in Section 5. Section 6 compares the conditional in-control (IC) and out-of-control (OOC) performance of the Hotelling T^2 and the SVDD-based charts. An real-life example is provided to show the SVDD-based chart's implementation in Section 7 and, finally, conclusions and future research directions complete the paper in Section 8.

2 The SVDD Algorithm

Assume that a Phase-I data set of size N contains a sequence of p-variate IC observations $\boldsymbol{x}_i = [x_{i1}, x_{i2}, \ldots, x_{ip}]^\intercal$, for $i = 1, 2, \ldots, N$. The goal of the SVDD algorithm is to find an optimal hyper-sphere with a minimum volume while capturing all possible training data. Let $\mathbf{O}$ be the center of the hyper-sphere and R be its radius. These quantities can be obtained by solving the following constrained optimization problem

$$\text{Min } R^2 + C\sum_{i=1}^{N}\xi_i, \tag{1}$$

$$\text{s.t. } \|\boldsymbol{x}_i - \mathbf{O}\|^2 \leq R^2 + \xi_i, \tag{2}$$

$$\xi_i \geq 0,\ i = 1, 2, \ldots, N, \tag{3}$$

where $\|\cdot\|$ represents the Euclidean distance, ξ_i, $i = 1, 2, \ldots, N$, are the so called *slack variables* and $C > 0$ is the penalty coefficient that controls the trade-off between the volume of the hyper-sphere and the misclassification errors.

For the optimization problem in (1), a Lagrangian function can be constructed as follows:

$$L = R^2 + C\sum_{i=1}^{N}\xi_i - \sum_{i=1}^{N}\eta_i(R^2 + \xi_i - \|\boldsymbol{x}_i - \mathbf{O}\|^2) - \sum_{i=1}^{N}\nu_i\xi_i, \tag{4}$$

where $\eta_i \geq 0$ and $\nu_i \geq 0$ are the Lagrange multipliers. Setting the partial derivatives of L with respect to R, $\mathbf{O}$ and ξ_i and equating them with zero, the following constraints are obtained:

$$\sum_{i=1}^{N}\eta_i = 1,\ \mathbf{O} = \sum_{i=1}^{N}\eta_i\boldsymbol{x}_i,\ C - \eta_i - \nu_i = 0,\ i = 1, 2, \ldots, N. \tag{5}$$

When substituting these constraints into (4), the optimization problem could be further simplified to

$$\text{Max } L = \sum_{i=1}^{N}\eta_i(\boldsymbol{x}_i \cdot \boldsymbol{x}_i) - \sum_{i=1,j=1}^{N}\eta_i\eta_j(\boldsymbol{x}_i \cdot \boldsymbol{x}_j), \tag{6}$$

$$\text{s.t. } \sum_{i=1}^{N}\eta_i = 1,\ 0 \leq \eta_i \leq C. \tag{7}$$

Solving this Quadratic Programming (QP) allows to obtain the optimal solution for the set $\boldsymbol{\eta} = \{\eta_1, \eta_2, \ldots, \eta_N\}$. Only the data points $\boldsymbol{x}_{SV}$ corresponding to non-zero η_i are needed in the description of the boundary, recorded as the Support Vector (SVs). We can find that, in (5), the center of the sphere $\mathbf{O}$ is expressed as a linear combination of the observations $\boldsymbol{x}_i$ with weights $\boldsymbol{\eta}_i$, $i = 1, \ldots, N$, and the squared-Euclidean distance (referred to as "distance" in the following) between $\boldsymbol{x}_{SV}$ and the center $\mathbf{O}$ is the square of the radius

$$R^2 = \|\boldsymbol{x}_{SV} - \mathbf{O}\|^2 = (\boldsymbol{x}_{SV} \cdot \boldsymbol{x}_{SV}) - 2\sum_{i=1}^{N}\eta_i(\boldsymbol{x}_{SV} \cdot \boldsymbol{x}_i) + \sum_{i=1,j=1}^{N}\eta_i\eta_j(\boldsymbol{x}_i \cdot \boldsymbol{x}_j). \tag{8}$$

A point $\mathbf{z}$ to be classified is declared to belong to the target category when the distance-function $df(\boldsymbol{z})$ is less than or equal to R^2, i.e.

$$df(\boldsymbol{z}) = (\boldsymbol{z} \cdot \boldsymbol{z}) - 2\sum_{i=1}^{N}\eta_i(\boldsymbol{z} \cdot \boldsymbol{x}_i) + \sum_{i=1,j=1}^{N}\eta_i\eta_j(\boldsymbol{x}_i \cdot \boldsymbol{x}_j) \leq R^2. \tag{9}$$

When the process data are linear inseparable, the kernel function method is applied to map the data into a high-dimensional feature space, so that the nonlinear problem in a low-dimensional space is transformed into the linear problem in a high-dimensional space. Replacing the inner product $(\boldsymbol{x}_i \cdot \boldsymbol{x}_j)$ with a kernel function $K(\boldsymbol{x}_i, \boldsymbol{x}_j)$ will lead to a more flexible decision boundary, and (8) and (9) become

$$R^2 = K(\boldsymbol{x}_{SV}, \boldsymbol{x}_{SV}) - 2\sum_{i=1}^{N}\eta_i K(\boldsymbol{x}_{SV}, \boldsymbol{x}_i) + \sum_{i=1,j=1}^{N}\eta_i\eta_j K(\boldsymbol{x}_i, \boldsymbol{x}_j), \tag{10}$$

$$df(\boldsymbol{z}) = K(\boldsymbol{z}, \boldsymbol{z}) - 2\sum_{i=1}^{N}\eta_i K(\boldsymbol{z}, \boldsymbol{x}_i) + \sum_{i=1,j=1}^{N}\eta_i\eta_j K(\boldsymbol{x}_i, \boldsymbol{x}_j). \tag{11}$$

For more details about the use of kernel functions, the interested reader can refer to Tax et al. (2004). In the following, a Gaussian kernel function $K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \exp(-\frac{\|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2}{s^2})$ is considered, where $s > 0$ is the kernel window width.

3 The SVDD-based control chart

Sun and Tsung (2003), among others, considered a SVDD-based control chart by plotting $df(\boldsymbol{z})$ as the monitoring statistics. They selected the parameters C and s to adjust the fraction of SVs to make it as close as possible to the Type-I error α, and subsequently, they set the R^2 as the control limit. However, their simulation results showed that this design can hardly achieve a satisfactory performance due to the limited parameter selection flexibility. In this paper, a separated threshold value h is set as the control limit and, if a point $\boldsymbol{z}$, satisfies $df(\boldsymbol{z}) > h$, then it is considered as being OOC.

Therefore, in the design of the SVDD-based control chart (if a Gaussian kernel function is considered), there are three major parameters to be set: the window width s; the penalty coefficient C; and the control limit h. Note that, the first two parameters, s and C, are internal parameters of the support vector methodology, while h is a counterpart of the Type-I error α. One may think that the addition of the parameter h will complicate the optimization. But as we will use the $(1-\alpha)$th quantile of the distance (instead of the R^2) of all training points to estimate the control limit h, there will be no additional optimization calculation concerning this parameter. Furthermore, this estimation of the control limit h also enhances the SVDD-based chart's robustness for the choice of different C values (this issue will be explained later).

For SVDD-based control charts, there is still no systematic method nor guideline for choosing the parameters, s and C, efficiently. The problem is that there is no knowledge about the classification error (i.e. no label for points incorrectly classified) to constructs a standard. Here, in order to better understand their effect, we intuitively select the appropriate parameters by describing the compactness and shape of the boundary for a two-dimensional data set. We generate a bivariate gamma-shape data set using the SAT model available in Cheng et al. (2017), see also equation (13). The first $N_I = 200$ samples are drawn from an IC process, representing the training/reference Phase-I data, and the remaining $N_{II} = 200$ samples represent the testing Phase-II data.

In Phase-I, Figure 1 compares the control boundaries of the SVDD obtained for several values $s \in \{1, 2, 4, 8, 16, 32\}$ when $C = 0.8$. We use "+" to represent the training samples, "•" to represent the SVs, and a solid line representing the data description boundary. It can be seen that s not only affects the radius, but it also affects the original shape of the boundary. A smaller value of s will make the description of the data more specific, it allows to obtain a closer description result, but this more flexible description boundary requires more SVs, for example, in Figure 1(a) when $s = 1$, almost all points are SVs. While, a larger value of s will make the shape of the boundary smoother, for example, in Figure 1(f) when $s = 32$, only two support vectors are required to completely determine the data description, but at the same time, we have to admit that this kind of boundaries is not representative and compact enough for describing the sample data distribution. The penalty coefficient, C, on the other hand, controls the number of points that fall outside the boundary. A larger value of C reduces the number of Phase-I points outside the boundary. As it appears in Figure 2, when C decreases, the number of SVs increases.

(Please insert Figures 1 and 2 here)

It is important to note that finding the optimal values of C and s using the reference sample without further information about the incorrect classifications is a difficult challenge. It is impossible to use the analytical method of minimizing the classification error rate. One can, however, get some

insights by visualizing the scatterplot for two or three-dimensional data. Therefore, this work use a two-dimensional data set as example. At first, we seek to find the appropriate values of s and C values to get a suitable boundary that can describe the Phase-I data the best as possible. After that, the value of h is estimated to set the Type-I error value at α. A detailed procedure is summarized as follows:

1. Concerning parameter s, it is found that a relatively large value of s makes the enveloping curve of the data more spherical and insensitive to the data shift, leading to more false alarms. Therefore, a smaller value of the parameter s should be recommended to obtain a boundary having a compact shape. For example, in the light of Figure 1, $s \in [4, 8]$ may be recommended.

2. Concerning parameter C, it can be seen that, as its value increases, the number of SVs tends to decrease, which is consistent with previous studies. Moreover, the $(1-\alpha)$th quantile of the distance of all training points N_I can provide a robust estimation of the control limit for a fixed α value regardless of the value of C.

3. Compute the distance of all training points $D_1 = df(\boldsymbol{x}_1), D_2 = df(\boldsymbol{x}_2), \ldots, D_N = df(\boldsymbol{x}_{N_I})$. Given a Type-I error α and the ordered D values $D_{(1)} < D_{(2)} < \cdots < D_{(N_I)}$, the $(1-\alpha)$th percentile of the ordered D values for all training N_I points is used to estimate the control limit $h = D_{(\lceil N_I \cdot (1-\alpha) \rceil)}$, where $\lceil \ldots \rceil$ denotes the rounded-up integer.

Through the above steps, for the Phase-II testing samples of size $N_{II} = 200$, Figure 3 draws the R^2 (using eq. (10)) and the control limit h values for different values of C when $\alpha = 0.05$ and $s = 4$. It can be seen that by taking different values of $C \in \{0.2, 0.6, 1.0\}$, the R^2 value is changing, while the value of h based on the quantile of the distance remains stable. This reduces the sensitivity of the control limit concerning the choice of C. So we suggest $C = 1$ in this paper, and thus the inequality constraint $0 \leq \eta_i \leq C$ in (7) can be relaxed. Therefore, we actually have only a single parameter s to search for with less optimisation complexity.

(Please insert Figure 3 here)

4 Conditional performance

It needs to be emphasised that although the SVDD-based chart pays little attention to how data are distributed, its performance strongly relies on the Phase-I training data set. Kumar et al. (2006) had considered the over-fitting problem with outliers existing in the Phase-I data set. In this paper, we will no further analyse this issue, but we will investigate the Phase-II performance of the SVDD-based chart when estimating the control limit h *conditional on* the training Phase-I samples.

The Average Run Length (ARL) is the most commonly used index to evaluate the performance of both parametric and nonparametric charts. The same $\text{ARL}_0 = 1/\alpha$ value is taken when the process is IC, and when the process is OOC, the smaller the $\text{ARL}_1 = 1/(1-\beta)$ value, the better the performance of control charts, where β is the Type-II error. However, as emphasized in Ning and Tsung (2013), it is impossible to set a too large value for ARL_0 (the in-control ARL) unless the number of Phase-I samples is very large. For example, if one wants to set $\text{ARL}_0 = 100$ then at least 1500 training samples are required! In addition, the SVDD algorithm also requires a high computational effort during the training stage. For example, it takes about 2.25 hours to train a procedure with just 3000 bivariate normal observations using Matlab on a Intel Core i7 CPU. Therefore, considering the speed of calculation and the ease of implementation, we have decided to only investigate the case

of $\text{ARL}_0 = 100$, as it is often used in SVDD-based charting literature literature (e.g. Kumar et al. (2006), Sukchotrat et al. (2009), Ning and Tsung (2013)).

In this paper, two bivariate models considering the dependence between variables are investigated:

- A symmetric one based on the bivariate normal $\text{BN}(a_X, b_X, a_Y, b_Y, \rho)$ distribution with probability density function defined as:

$$f_{\text{BN}}(x, y|a_X, b_X, a_Y, b_Y, \rho) = \left(2\pi b_X b_Y \sqrt{1-\rho^2}\right)^{-1} \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{(y-a_X)^2}{b_Y^2} - \frac{2\rho(y-a_X)(x-a_Y)}{b_X b_Y} + \frac{(x-a_Y)^2}{b_X^2}\right)\right], \tag{12}$$

  where a_X and $b_X > 0$ (a_Y and $b_Y > 0$) are the mean and standard deviation parameters of X and Y, respectively, and $\rho \in (0, 1)$ is the correlation coefficient.

- An asymmetric one based on the bivariate gamma $\text{SAT}(a_X, b_X, a_Y, b_Y, \rho)$ distribution with probability density function defined for $x > 0$, $y > 0$ as:

$$f_{\text{SAT}}(x, y > 0|a_X, b_X, a_Y, b_Y, \rho) = \frac{\left(\frac{x}{b_X}\right)^{a_X - 1}\left(\frac{y}{b_Y}\right)^{a_Y - 1} \exp\left(-\frac{\left(\frac{x}{b_X}\right)+\left(\frac{y}{b_Y}\right)}{1-\rho\sqrt{a_Y/a_X}}\right)}{\left(1-\rho\sqrt{a_Y/a_X}\right)^{a_X} \Gamma(a_X)\,\Gamma(a_Y - a_X)} \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \frac{\left(\rho\sqrt{a_Y/a_X}\right)^{j+k} \Gamma(a_Y - a_X + k)}{\left(1-\rho\sqrt{a_Y/a_X}\right)^{2j+k} \Gamma(a_Y + j + k)\, j!\, k!} \left(\frac{x}{b_X}\right)^{j} \left(\frac{y}{b_Y}\right)^{j+k}, \tag{13}$$

  where $a_X > 0$ and $b_X > 0$ ($a_Y \geq a_X$ and $b_Y > 0$) are the corresponding shape and scale parameters of Y and X, respectively, $\Gamma(\cdot)$ is the gamma function, and $\rho \in (0, 1)$ is a parameter that quantifies the correlation.

In order to have an overview of the conditional IC performances of the SVDD-based chart, Table 1 presents the percentiles of the IC ARL_0 values based on several numbers of Phase-I training samples $N_I \in \{50, 200, 500, 1000, 2000, 3000\}$. The targeted IC ARL_0 value is set to 100 (i.e. $\alpha = 0.01$). According to the explanations provided in Section 3, a value of $s = 8$ is suggested for the SVDD-based chart. We suggest the following steps for obtaining the IC ARL_0's:

1. Generate N_I training observations $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_{N_I}$ from IC processes and compute the distance of all training points $D_1 = df(\boldsymbol{x}_1), D_2 = df(\boldsymbol{x}_2), \ldots, D_N = df(\boldsymbol{x}_{N_I})$. Given a Type-I error α, the control limit $h = D_{(\lceil N_I \cdot (1-\alpha) \rceil)}$ is estimated as in Section 3.

2. Generate $N_{II} = 5000$ Phase-II IC testing observations $\boldsymbol{z}_1, \boldsymbol{z}_2, \cdots, \boldsymbol{z}_{N_{II}}$, and then compute the statistics $df(\boldsymbol{z}_i)$, $i = 1, 2, \ldots, N_{II}$. Record the run length $\text{RL} = i$ if a first signal is observed $df(\boldsymbol{z}_i) > h$. If no OOC signal is raised, i.e. $i > N_{II}$ then the value $\text{RL} = N_{II} = 5000$ will be recorded.

3. Repeat step 2 (10^4 times), and then average the run lengths IC ARL_0.

4. Repeat steps 1-3 (10^4 times) in order to obtain the IC ARL_0 empirical distribution.

Observations in Table 1 show that the ARL_0 values of the SVDD-based chart vary a lot as the number of Phase-I samples changes. When a too small value of N_I is selected (e.g., 50, 200), the ARL_0 of the SVDD-based charts may not reach the target value. For example, if we use a training sample size $N_I = 50$ for the BN$(10, 1, 10, 2, 0.5)$ distribution, 95% of the practitioners would have an ARL_0 value less than 46.2. This is far too less from the target value of 100. This happens because the SVDD method is completely data-driven and a small Phase-I sample size cannot describe the data distribution well. In this case, no matter the value of the target ARL_0, the attained ARL_0 is always less. The same conclusion can also be found in Ning and Tsung (2013). As N_I increases up to 500 or 1000, the 50^{th} quantile of ARL_0 values, i.e. the median of ARL_0 ($MARL_0$), gets closer to 100. Still, in more than 50% cases, the attained ARL_0 value will be below the target one, and these charts are also associated with a large variation in the ARL_0 distribution. The standard deviations of the ARL_0 ($SDARL_0$) are recorded in the last columns. A smaller $SDARL_0$ will provide more confidence to the practitioners about the charts' actual IC performance. For example, if we have $N_I = 2000$ and 3000 for the BN$(10, 1, 10, 2, 0.5)$ distribution, the $SDARL_0$ values decrease to 21.6 and 16.8 , respectively. As suggested in Zhang et al. (2014), a $SDARL_0$ within 10% of the of ARL_0 may be reasonable. Therefore, only a larger number of Phase-I samples may guarantee that individual practitioners can obtain an ARL_0 value close to the target value.

(Please insert Table 1 here)

5 Adjusting the control limits

The SVDD-based chart can detect shifts without knowing the underlying distribution, but it requires many training samples to ensure a reasonable IC robustness for different Phase-I data sets. In the design of parametric control charts with estimated parameters, Saleh et al. (2015), Hu et al. (2018), Tang et al. (2019), among others, suggested the use of bootstrap-type algorithms to adjust the control limits to guarantee that the conditional probability of the ARL_0 value to exceed the expected value will be close to a pre-specified constant, say $(1 - \epsilon)100\%$ (e.g. $\epsilon = 0.1$). In this paper, two bootstrap methods using the Phase-I data are discussed to overcome the problem of low attained ARL_0 due to the "between-practitioners variability".

The bootstrap approach helps in constructing nonparametric confidence intervals for the control limit by re-sampling the Phase-I data. Both bootstrap percentile and bootstrap percentile-t methodologies are used in this paper. Of course, other types of bootstrap methods, like the bootstrap BC_a method or the additively corrected bootstrap-t, can also be applied to solve this problem, see Davison and Hinkley (1997) or Polansky (2000). We studied their performance and observed that as the number of Phase-I samples increases, all these methods provide nearly the same results. They only differ when $N_I \leq 500$. Therefore, for brevity, we only include in this paper two typical types of bootstrap methods. A more explicit description of the two bootstrap procedures used in this paper are given below:

Bootstrap Percentile

1. Generate $l = 1, 2, \ldots, B$ bootstrap samples $\boldsymbol{x}_{l1}, \boldsymbol{x}_{l2}, \ldots, \boldsymbol{x}_{lN_I}$ of size N_I from the Phase-I data set and compute the corresponding distance of all points $D_{l1} = df(\boldsymbol{x}_{l1}), D_{l2} = df(\boldsymbol{x}_{l2}), \ldots, D_{lN_I} = df(\boldsymbol{x}_{lN_I})$ as a sequence of statistics from the lth bootstrap sample.

2. For each bootstrap sample l, compute the $(1-\alpha)$ percentile $D^*_{l(\lceil N_I \cdot (1-\alpha) \rceil)}$ of the ordered values $D^*_{l(1)} < D^*_{l(2)} < \cdots < D^*_{l(N_I)}$, which represents the estimated control limit $\hat{h}^*_l$.

3. Obtain the $(1-\epsilon)$ percentile of the bootstrap distribution $\hat{h}_l^*$, $l = 1, 2, \ldots, B$ as the adjusted control limit $h = \hat{h}^*_{(\lceil B\cdot(1-\epsilon)\rceil)}$.

Bootstrap Percentile-t

1. For original data set $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{N_I}$, compute the $(1-\alpha)$ percentile $D_{(\lceil N_I\cdot(1-\alpha)\rceil)}$ of the Phase-I observations as an estimated value of the control limit $\hat{h}$.

2. Generate $l = 1, 2, \ldots, B$ bootstrap samples of size N_I from the Phase-I data set, and compute the corresponding $\hat{h}_l^* = D^*_{l(\lceil N_I\cdot(1-\alpha)\rceil)}$ and $S_{\hat{h}_l^*}$, where $S_{\hat{h}_l^*}$ is an estimate of the standard error of $\hat{h}_l^*$ based on the sample l. The calculation of $S_{\hat{h}_l^*}$(and $S_{\hat{h}}$) can also be based on a bootstrap estimate.

3. For each bootstrap sample, compute the $\hat{t}_l^* = (\hat{h}_l^* - \hat{h})/S_{\hat{h}_l^*}$, $l = 1, 2, \ldots, B$.

4. Obtain the ϵ percentile of the bootstrap distribution $\hat{t}_l^*$, $l = 1, 2, \ldots, B$, and compute the adjusted control limit $h = \hat{h} - \hat{t}^*_{(\lceil B\cdot\epsilon\rceil)} \cdot S_{\hat{h}}$.

Figures 4 and 5 display the box plots of the IC ARL_0 distributions for the SVDD-based charts for different Phase-I sample sizes, when $B = 1000$, $\alpha = 0.01$ and $\epsilon = 0.1$. For brevity, we consider $N_I \in \{500, 1000, 1500, 2000\}$. The values calculated via the two types of bootstrap adjustments are referred as to 'Adjusted-p' and 'Adjusted-t', respectively. Otherwise, they are referred as to 'Unadjusted'.

For example, Figure 4 shows that for the $\text{BN}(10, 1, 10, 2, 0.5)$ observations, when $N_I = 500$, more than 50% of the SVDD-based charts with unadjusted control limits would have an IC ARL_0 below 100. On the other hand, as expected, the results with adjusted limits, have a higher level of MARL_0, but this is also not enough to ensure a reasonable variation in the ARL_0 values. Comparatively, the bootstrap percentile-t method is more appropriate than the bootstrap percentile method in this case. When N_I increases up to 1000, the 'Adjusted' cases result in more than 75% of the SVDD-based charts having an ARL_0 of 100 or more. Moreover, it can be seen that increasing the Phase-I training sample size N_I reduces the variation in SDARL_0 in Phase-II. At the same time, we observe that, the bootstrap percentile method has a slightly smaller SDARL_0 value than the value of the bootstrap percentile-t method when $1000 \leq N_I \leq 2000$. Therefore, the bootstrap percentile method provides a better conditional performance when the number of Phase-I samples is larger. Figure 5 exhibits similar trends as for Figure 4 but for the $\text{SAT}(4, 2.5, 100, 0.1, 0.5)$ distribution in place of the $\text{BN}(10, 1, 10, 2, 0.5)$ distribution.

From these results, clearly, the SVDD-based chart will require more Phase-I sample observations than parametric charts with estimated parameters. But we need to emphasize that, for the parametric charts, when the distributional assumptions are violated, it leads to unsatisfactory IC performances and also the control limit adjustments based on the parametric bootstrap method are no longer appropriate. Therefore, the SVDD-based chart could be an excellent practical alternative in such situations since it is completely data-driven.

(Please insert Figures 4 and 5 here)

6 Performance comparison

In this Section, a study is conducted to compare the conditional performance between the SVDD-based chart and and the Hotelling T^2 chart. It is important to note that the aim of this section is not to show the superiority of the SVDD-based chart over conventional charts but to complete the comparative works in the case of estimated parameters conditional on the Phase I sample. As mentioned in Sun and Tsung (2003), the relationship between the conventional charts and the SVDD-based chart is not 'one replaces the other' or 'one is better than the other', but it is rather 'one complements the other'. The comparisons of these charts are are readily available when assuming that the true process parameters are known.

6.1 Evaluation of IC robustness

The same data distributions considered in the previous Section are used to compare the performance of all charts, and the IC mean vector $\boldsymbol{\mu_0}$ and the IC covariance matrix $\boldsymbol{\Sigma_0}$ are listed below,

- for the bivariate normal BN$(10, 1, 10, 2, 0.5)$:
  $\boldsymbol{\mu_0} = \begin{bmatrix}\mu_X \\ \mu_Y\end{bmatrix} = \begin{bmatrix}10 \\ 10\end{bmatrix} \quad \boldsymbol{\Sigma_0} = \begin{bmatrix}\sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2\end{bmatrix} = \begin{bmatrix}1 & 1 \\ 1 & 4\end{bmatrix},$
- for the bivariate gamma SAT$(4, 2.5, 100, 0.1, 0.5)$:
  $\boldsymbol{\mu_0} = \begin{bmatrix}\mu_X \\ \mu_Y\end{bmatrix} = \begin{bmatrix}10 \\ 10\end{bmatrix} \quad \boldsymbol{\Sigma_0} = \begin{bmatrix}\sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2\end{bmatrix} = \begin{bmatrix}25 & 2.5 \\ 2.5 & 1\end{bmatrix}.$

Our task is to test the hypotheses $\{H_0 : \boldsymbol{\mu} = \boldsymbol{\mu_0}$ versus $H_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu_0}\}$. Here, for each distribution, we generate a data set consisting of $N_I = 2000$ IC Phase-I training observations. The control limits of all parametric charts are designed with $\text{ARL}_0 = 100$ ($\alpha = 0.01$) when the underlying distributions, as well as the parameters, are known, and we also adjust their control limits so that they produce a conditional IC ARL_0 that exceeds the targeted value with a certain probability with $(1 - \epsilon)100\%$ (e.g. $\epsilon = 0.1$). All results have been estimated via Monte Carlo simulation (10^4 runs).

Figure 6 presents a comparison of the box plots of the ARL_0 values for both specified distributions. As expected, the Hotelling T^2 chart has a good IC performance when the actual distribution is the bivariate normal one. On the other hand, under the bivariate gamma distribution, the Hotelling T^2 chart designed for the normal distribution is highly sensitive to normality violations, as the ARL_0 values are much smaller than the designed one. Human et al. (2011) mentioned that the IC robustness is the key to the proper design and implementation of any control charts. If charts are not IC robust, their shift detection capability in the OOC situation is almost meaningless. Clearly, in this example, the parametric charts' IC properties are greatly affected by a change in the underlying distribution and parameter estimation. So it should be considered with caution for quality practitioners to use parametric charts in such situations.

On the other hand, the SVDD-based chart has a stable IC performance for both distributions under consideration. However, it is also important to note that the unadjusted control limit can hardly guarantee the IC performance of the SVDD-based chart. The median ARL_0 is very close to the desired 100 when $N_I = 2000$, but there is a wide "between-practitioners variability" in the attained ARL_0. In this case, more than 50% of the charts have an IC ARL_0 value below 100. Therefore, we suggest adjusting the control limit based on the nonparametric bootstrap method proposed in Section 4 to guarantee a minimum IC performance with a prespecified probability.

(Please insert Figure 6 here)

6.2 Performance in signal detection

When the process is OOC, we are interested in detecting shifts in the process mean $\boldsymbol{\mu_1} = \boldsymbol{\mu_0} + \boldsymbol{\delta}$ from the nominal value, assuming that the IC covariance matrix does not change. Here, for each distribution defined in Section 4, a one standard-deviation shift of size $\boldsymbol{\delta} = [\sigma_X, \sigma_Y]^\intercal$ in the process mean has occurred, which means that the OOC mean value is $\boldsymbol{\mu_1} = [11, 12]^\intercal$ for the bivariate normal distribution, and is $\boldsymbol{\mu_1} = [15, 11]^\intercal$ for the bivariate gamma distribution.

The control limits of the Hotelling T^2 chart are redesigned so that it produces the same IC ARL_0 performance for a fair comparison of the OOC properties. From Figure 7, as it can be observed, the SVDD-based chart has a similar MARL_1 performance as the Hotelling T^2 chart when the underlying process follows the bivariate normal distribution. It is also noted that the SVDD-based chart has a slightly larger SDARL_0 value. For the bivariate gamma distribution, the Hotelling T^2 chart has more larger ARL_1 values, while the SVDD-based chart provides a better ARL_1 performance than its parametric counterparts.

It can be concluded that the Hotelling T^2 chart, which is designed for the multivariate normal distribution, could potentially be (highly) affected when the distributional assumption is violated. Although bootstrapping the Phase I sample may be performed in a nonparametric way, it requires some distributional assumption in order to guarantee the IC performance for small Phase-I samples. On the other hand, the SVDD-based chart could be an good choice since it has an acceptable IC robustness and a comparable shift detection capability.

(Please insert Figure 7 here)

7 An illustrative example

The real data set used in this section is freely accessible at the UC Irvine Machine Learning Repository (http://archive.ics.uci.edu/ml/datasets/SECOM) and it comprises 1567 observations (sometimes missing) on 591 variables collected from sensors of a complex modern semi-conductor manufacturing process. In the following illustrative example we will only focus on variables #31 and #32 for which 1463 IC observations are available. Figure 8(a) presents the scatter-plot for the 1463 observations of variables #31 and #32. In order to test the bivariate normality for these variables, we simply performed a Jarque-Bera test and the results clearly showed that their respective marginal distributions cannot be approximated with a normal distribution. This led us to conclude that the bivariate normality cannot be assumed for variables #31 and #32 and we could expect that the SVDD-based chart would be more robust than the Hotelling T^2 chart for monitoring this process.

To illustrate the use of the SVDD-based chart, we used the first 1200 samples as Phase-I training data, and then monitored the remaining 263 samples for evaluating the conditional performance. As discussed in Section 3, since $C = 1$ is fixed, the only parameter to be determined is s. The choice of s influences how well the boundary fits the data distribution. Thus, observing Figure 8(a), a value of $s = 8$ is suggested. Because the SVDD model is established from a small reference sample, a more proper procedure is to use an adjusted control limit based on the bootstrap percentile method described in Section 4. When the desired Type-I error $\alpha = 0.01$ ($\text{ARL}_0 = 100$) and $\epsilon = 0.1$ values are fixed, for each bootstrap sample $l = 1, 2, \ldots, 1000$, the $(1-\alpha)$ percentile $df^*_{l(\lceil N\cdot(1-\alpha)\rceil)}(\boldsymbol{x})$ is computed as the estimated control limit $\hat{h}^*_l$. Then finally, the $(1-\epsilon)$ percentile of the bootstrap distribution $\hat{h}^*_l$, $l = 1, 2, \ldots, B$ is obtained as the adjusted control limit $h = \hat{h}^*_{(\lceil B\cdot(1-\epsilon)\rceil)} = 0.53$. The Phase-II data

of the statistic $df(\boldsymbol{z})$, together with the adjusted control limit h, are also plotted in Figure 8(b). We can see that the SVDD-based chart detect an OOC situation at the 186 time points.

(Please insert Figure 8 here)

8 Concluding remarks

For a long time, the important effect of different Phase-I data on nonparametric-type control charts has been neglected. In this paper, we evaluated the conditional effect of the Phase-I samples on the Phase-II efficiency of a nonparametric chart based on the SVDD theory. The results indicated that many Phase-I samples are necessary to guarantee a high probability of the IC performance to be close to the desired target. Moreover, to take the "between-practitioners variability" into consideration, two bootstrap-based approaches have been proposed to estimate the control limit in this study. Although the SVDD-based chart usually needs somewhat more Phase-I samples, they prevent performance deterioration when the underlying distributional assumption is violated. Comparative results enable us to conclude that the SVDD-based chart is a good alternative when the distribution is unknown.

We encourage for more future researches on the optimal data-dependent choice of C and s. Finally, it may be worth applying cluster analysis on the Phase-I data to partition the IC data into few clusters and to identify support vector-based region for individual clusters and consider the combined region as an in-control one. However, determining the optimal number of clusters and the corresponding combining strategies would be more complex and it is left for future researches.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Alevizakos V, Koukouvinos C and Chatterjee K. A nonparametric double generally weighted moving average signed-rank control chart for monitoring process location. *Quality and Reliability Engineering International*, 2020, 36(7): 2441–2458.

Asghari S, Gildeh B, Ahmadi J and Borzadaran G. Sign control chart based on ranked set sampling. *Quality Technology & Quantitative Management*, 2018, 15(5): 568–588.

Castagliola P and Maravelakis P. A CUSUM control chart for monitoring the variance when parameters are estimated. *Journal of Statal Planning & Inference*, 2011, 141(4): 1463–1478.

Castagliola P, Tran K, Celano G, Rakitzis A and Maravelakis P. An EWMA-Type sign chart with exact run length properties. *Journal of Quality Technology*, 2019, 51(1), 51–63.

Chakraborti S and Graham M. Nonparametric distribution-free control charts: An updated overview and some results. *Quality Engineering*, 2019, 31(4): 523–544.

Chen K, Chang T, Wang K, and Huang C. Developing control charts in monitoring service quality based on the number of customer complaints. *Total Quality Management & Business Excellence*, 2015, 26(5): 675–689.

Cheng Y, Mukherjee A and Xie M. Simultaneously monitoring frequency and magnitude of events based on bivariate gamma distribution. *Journal of Statistical Computation and Simulation*, 2017, 87(9): 1723–1741.

Cheng X and Wang F. VSSI median control chart with estimated parameters and measurement errors. *Quality & Reliability Engineering International*, 2018, 34(5): 867–881.

Davison A and Hinkley D. Bootstrap methods and their application. *Cambridge University Press*, 1997.

Diko, Goedhart R, Chakraborti S, Does R and Epprecht E. Phase II control charts for monitoring dispersion when parameters are estimated. *Quality Engineering*, 2017, 29(4): 605–622.

Dovoedo Y and Chakraborti S. Effects of parameter estimation on the multivariate distribution-free Phase II sign EWMA Chart. *Quality and Reliability Engineering International*, 2016, 33(2): 431–449.

Gani W, Taleb H and Limam M. An assessment of the kernel-distance-based multivariate control chart through an industrial application. *Quality and Reliability Engineering International*, 2011, 27(4): 391–401.

Graham M, Mukherjee A and Chakraborti S. Design and implementation issues for a class of distribution-free Phase-II EWMA exceedance control charts. *International Journal of Production Research*, 2017, 55(8): 2397–2430.

Holland M and Hawkins D. A control chart based on a nonparametric multivariate change-point model. *Journal of Quality Technology*, 2014, 46(1), 63–77.

Hu X, Castagliola P, Ma Y and Huang W. Guaranteed in-control performance of the synthetic chart with estimated parameters. *Quality and Reliability Engineering International*, 2018, 34(5): 759–771.

Human S, Kritzinger P and Chakraborti S. Robustness of the EWMA control chart for individual observations. *Journal of Applied Statistics*, 2011, 38(10): 2071–2087.

Huwang L, Lin L, and Yu C. A spatial rank-based multivariate EWMA chart for monitoring process shape matrices. *Quality and Reliability Engineering International*, 2019, 35(6): 1716–1734.

Kumar S, Choudhary A, Kumar M, Shankar R and Tiwari M. Kernel distance-based robust support vector methods and its application in developing a robust K-chart. *International Journal of Production Research*, 2006, 44(1): 77–96.

Mahmoud M and Maravelakis P. The Performance of multivariate CUSUM control charts with estimated parameters. *Journal of Statistical Computation and Simulation*, 2013, 83(4): 721-šC738.

Ning X and Tsung F. Improved design of kernel distance-based charts using support vector methods. *IIE Transactions*, 2013, 45(4): 464–476.

Polansky A. Stabilizing bootstrap-t confidence intervals for small samples. *Canadian Journal of Statistics*, 2000, 28(3):501–516.

Qiu P. Some perspectives on nonparametric statistical process control. *Journal of Quality Technology*, 2018, 50(1): 49–65.

Saleh N, Mahmoud M, Jones-Farmer L, Zwetsloot I and Woodall W. Another look at the EWMA control charts with estimated parameters. *Journal of Quality Technology*, 2015, 47(4): 363–382.

Singh A and Prasher A. Measuring healthcare service quality from patients' perspective: using fuzzy AHP application. *Total Quality Management & Business Excellence*, 2019, 30: 284–300.

Smith O, Adelfang S and Tubbs J. A bivariate gamma probability distribution with application to gust modeling. *NASA Technical Memorandum 82483*, 1982.

Sukchotrat T, Kim S and Tsung F. One-class classification-based control charts for multivariate process monitoring. *IIE Transactions*, 2009, 42(2): 107–120.

Sun R and Tsung F. A kernel-distance-based multivariate control chart using support vector methods. *International Journal of Production Research*, 2003, 41(13): 2975–2989.

Tang A, Sun J, Hu X and Castagliola P. A new nonparametric adaptive EWMA control chart with exact run length properties. *Computers & Industrial Engineering*, 2019, 130: 404–419.

Tang A, Castagliola P, Hu X and Sun J. The adaptive EWMA median chart for known and estimated parameters. *Journal of statistical computation and simulation*, 2019, 89(5): 844–863.

Tanveer H, Khan M, Salman K, Amin U, Mi Y and Sung W. Intelligent baby behavior monitoring using embedded vision in IoT for smart healthcare centers. *Ai & Society*, 2019, 1(1): 110–124.

Tax D and Duin R. Support vector data description. *Machine learning*, 2004, 54(1): 45–66.

Woodall W. The use of control charts in health-care and public-health surveillance. *Journal of Quality Technology*, 2006, 38(2): 89–104.

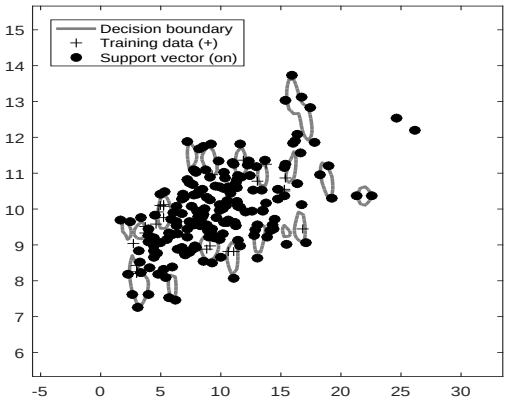
Zhang M, Megahed F and Woodall W. Exponential CUSUM charts with estimated control limits. *Quality and Reliability Engineering International*, 2014, 30(2):275–286.

Zou C and Tsung F. A multivariate sign EWMA control chart. *Technometrics*, 2011, 53(1): 84-97.

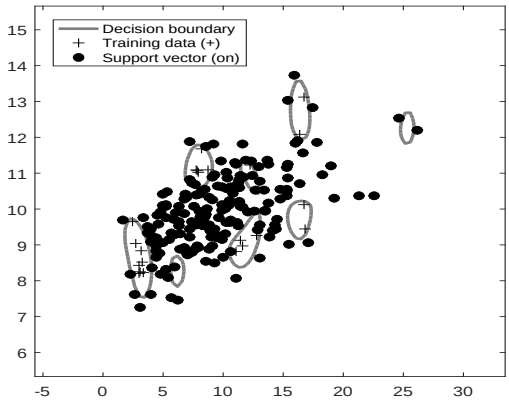
Zou C, Wang Z and Tsung F. A spatial rank-based multivariate EWMA control chart. *Naval Research Logistics*, 2012, 59(2): 91–110.

Table 1: Percentiles of the IC ARL_0 values based on several numbers of training Phase-I samples

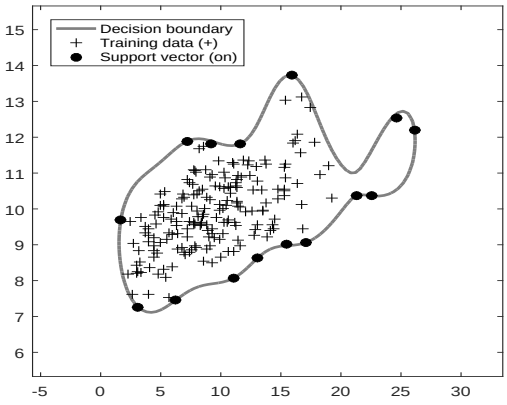
F	N_I	5th	10th	25th	50th	75th	90th	95th	AARL	SDARL
BN	50	9.7	11.5	15.5	21.6	29.9	39.4	46.2	24.2	12.6
	200	30.4	34.7	42.6	54.3	71.4	95.2	111.1	60.6	27.5
	500	51.8	58.5	72.2	90.9	120.5	153.8	185.2	99.4	45.9
	1000	60.2	64.9	75.8	91.7	116.4	137.0	155.0	99.5	30.8
	2000	71.9	76.9	87.0	99.0	114.9	129.8	140.8	100.9	21.6
	3000	77.5	81.3	89.3	100.1	112.4	123.5	129.9	101.0	17.0
SAT	50	7.2	7.9	9.0	12.8	16.8	22.5	26.8	14.3	6.8
	200	21.4	24.1	31.2	38.9	51.3	67.8	86.2	44.2	22.1
	500	45.7	49.8	61.5	82.6	107.5	140.8	161.3	89.5	37.6
	1000	59.2	65.4	78.1	92.2	121.9	151.5	166.7	98.9	36.4
	2000	70.4	76.0	87.0	100.5	119.0	138.9	153.8	100.4	26.1
	3000	72.5	76.9	89.0	100.6	104.2	116.3	123.5	100.3	16.7



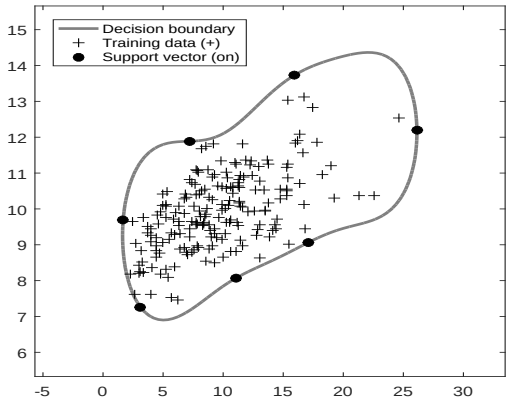
(a) $C = 0.8$, $s = 1$



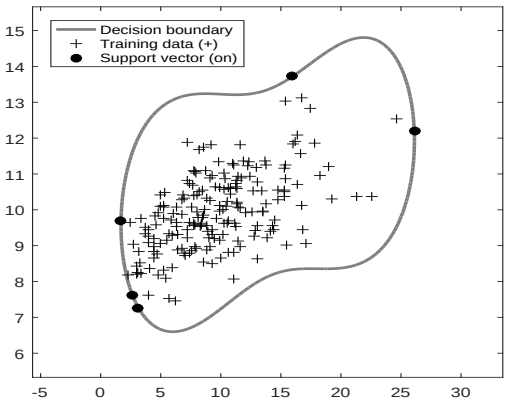
(b) $C = 0.8$, $s = 2$



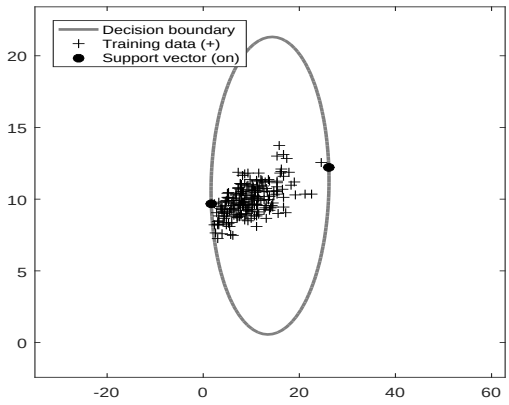
(c) $C = 0.8$, $s = 4$



(d) $C = 0.8$, $s = 8$

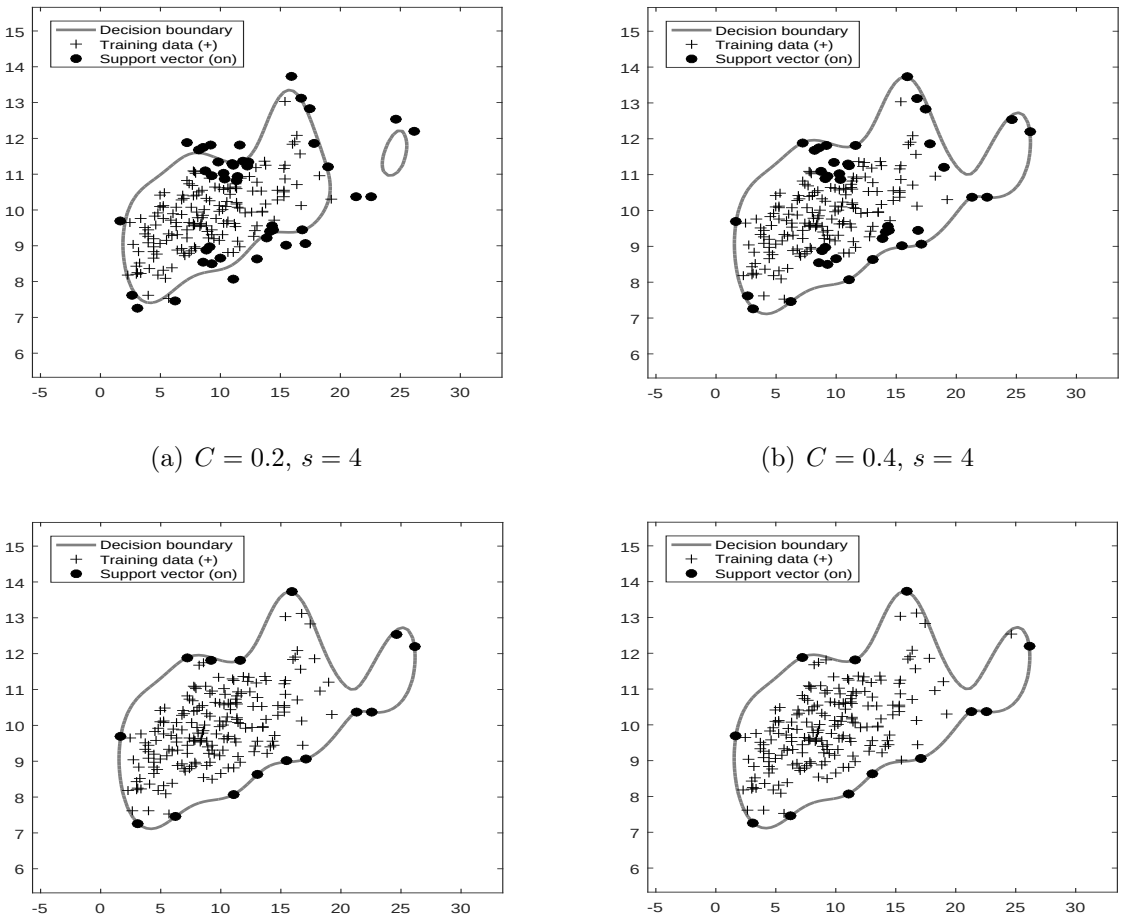


(e) $C = 0.8$, $s = 16$



(f) $C = 0.8$, $s = 32$

Figure 1: Control boundaries of the SVDD obtained for different values of s when $C = 0.8$ and $N = 200$.



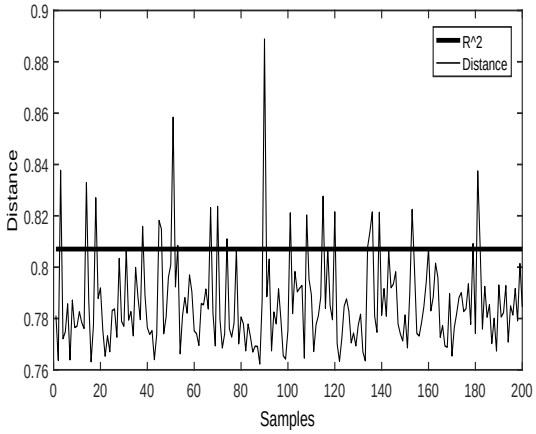
(a) $C = 0.2$, $s = 4$

(b) $C = 0.4$, $s = 4$

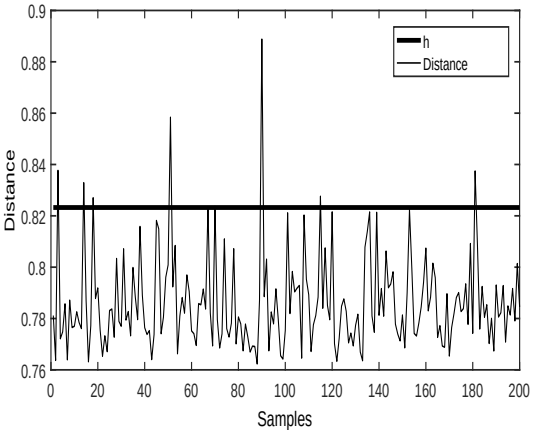
(c) $C = 0.8$, $s = 4$

(d) $C = 1.0$, $s = 4$

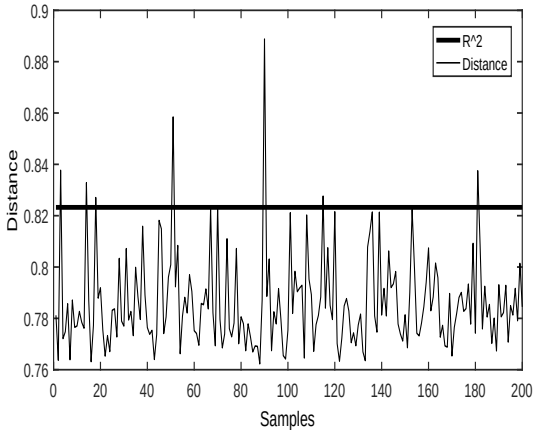
Figure 2: Control boundaries of the SVDD obtained for different values of C when $s = 4$ and $N = 200$.



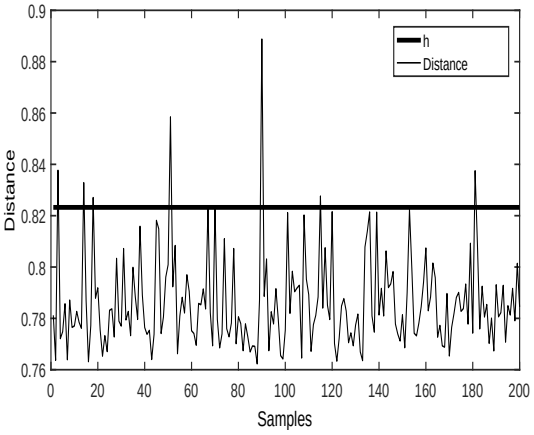
(a) The R^2 value for $C = 0.2$



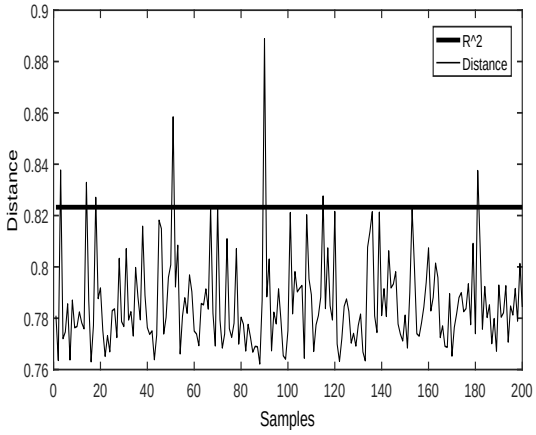
(b) The h value for $C = 0.2$



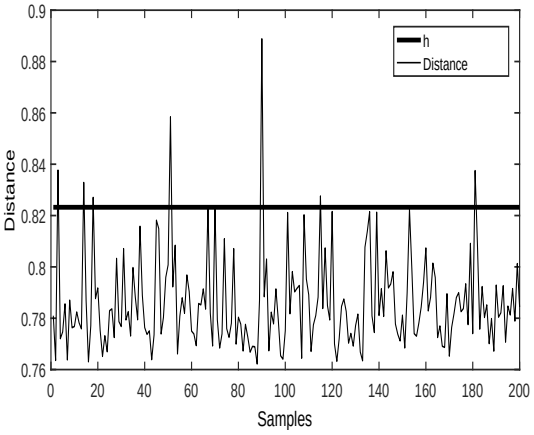
(c) The R^2 value $C = 0.6$



(d) The h value for $C = 0.6$

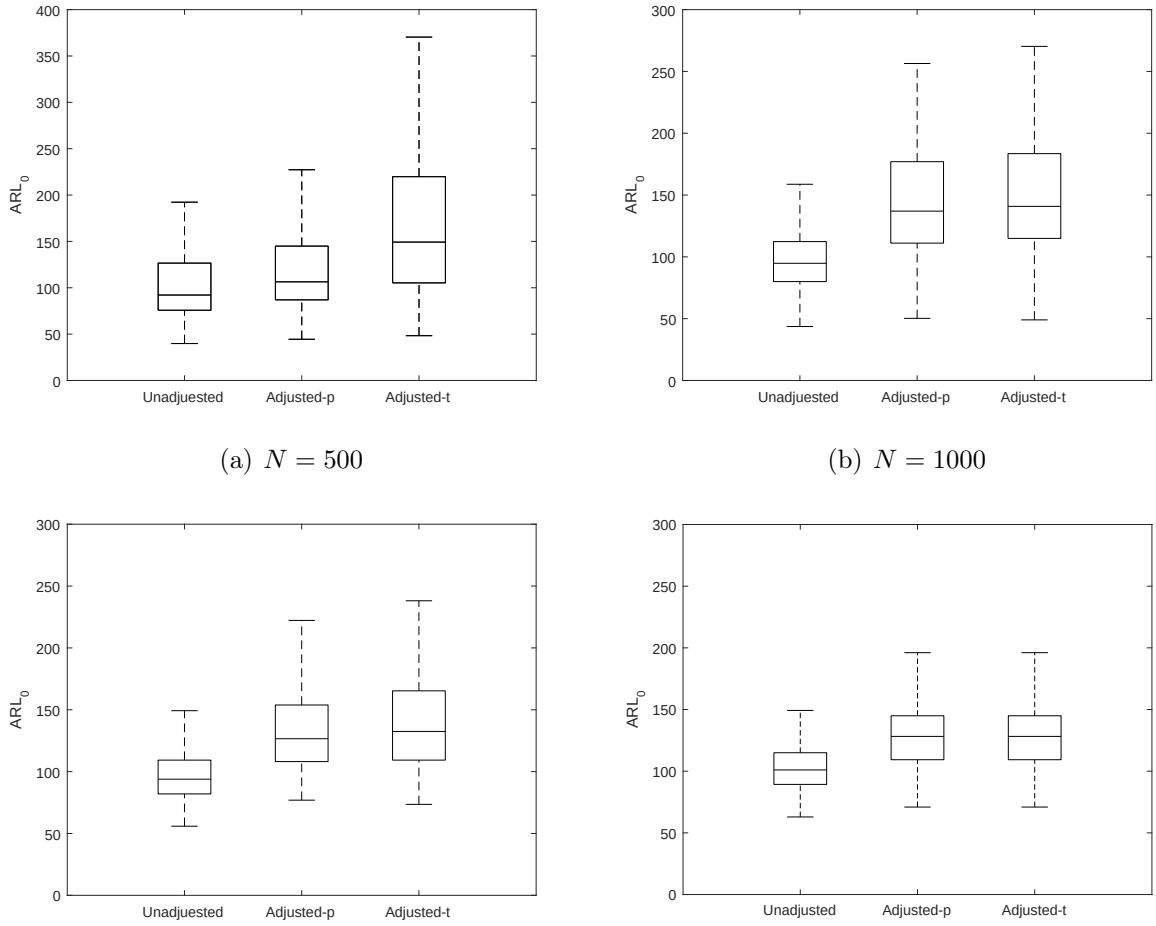


(e) The R^2 value for $C = 1.0$



(f) The h value for $C = 1.0$

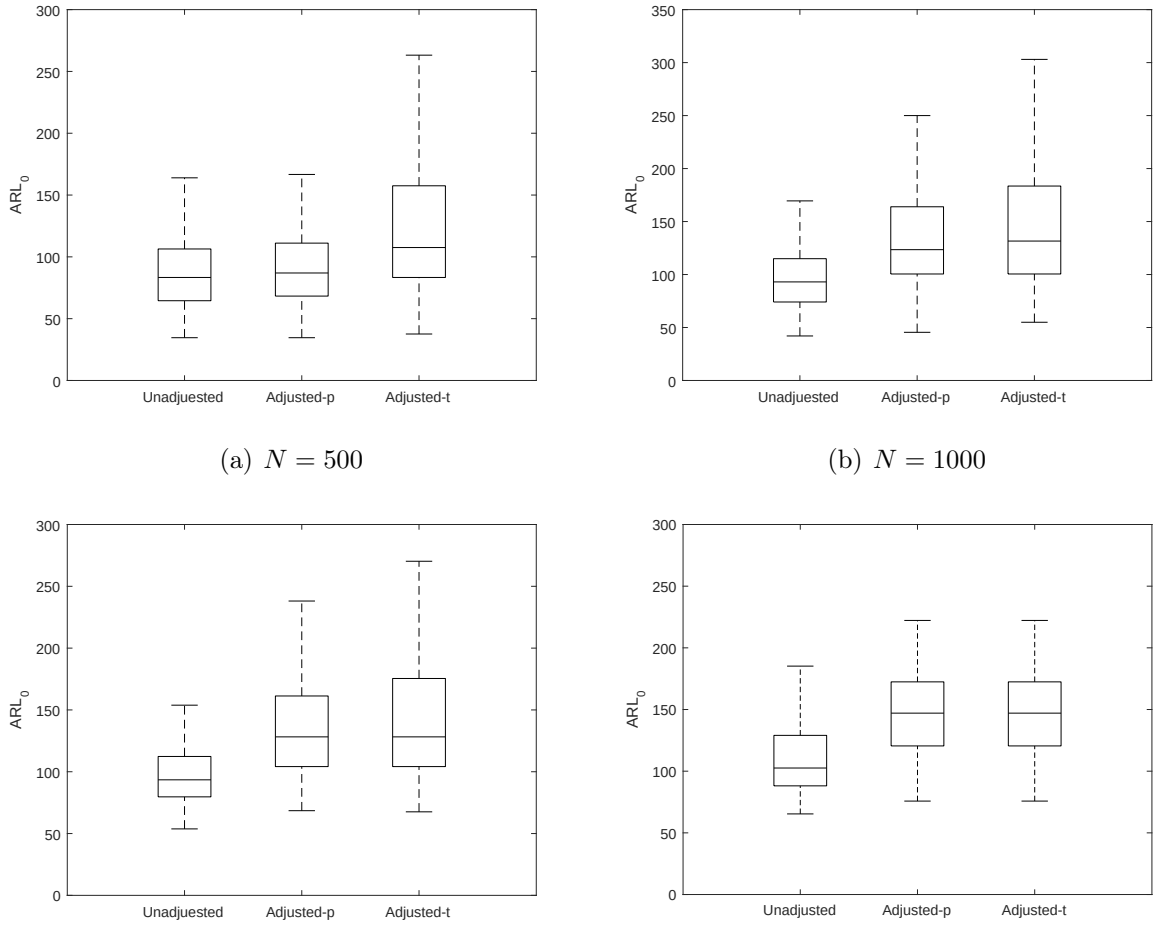
Figure 3: The R^2 and the control limit h values of the SVDD obtained for different values of C when $s = 4$, $\alpha = 0.05$ and $N = 200$.



(a) $N = 500$ (b) $N = 1000$

(c) $N = 1500$ (d) $N = 2000$

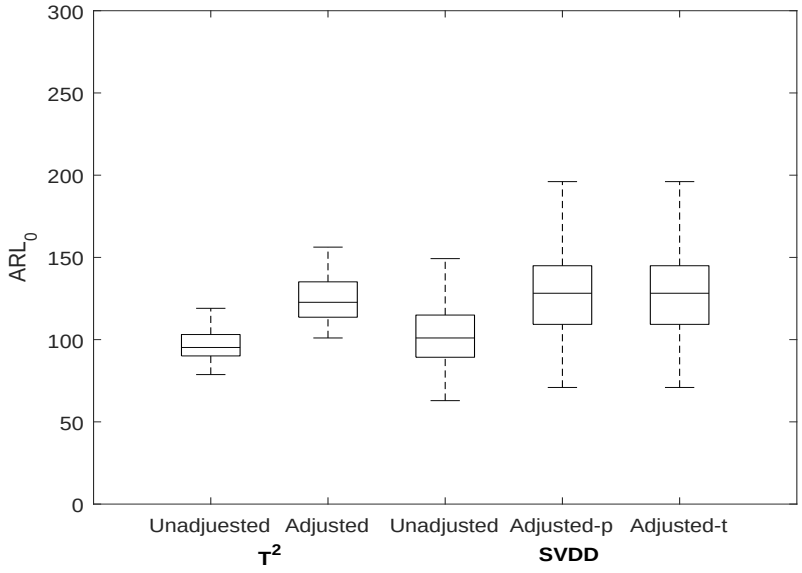
Figure 4: The box plots of the IC ARL_0 with the unadjusted and adjusted limits for the $\text{BN}(10, 1, 10, 2, 0.5)$ distribution when $s = 8$, $B = 1000$ and $\alpha = 0.01$.



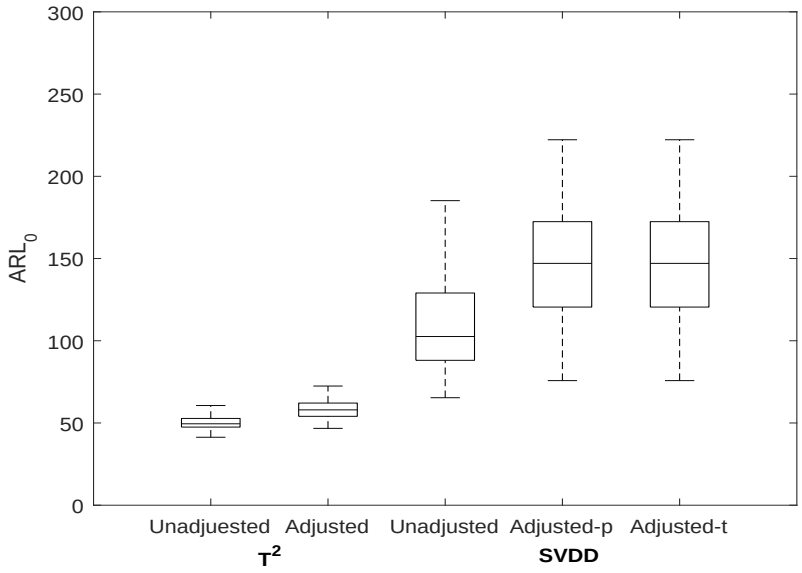
(a) $N = 500$ (b) $N = 1000$

(c) $N = 1500$ (d) $N = 2000$

Figure 5: The box plots of the IC ARL_0 with the unadjusted and adjusted limits for the $\text{SAT}(4, 2.5, 100, 0.1, 0.5)$ distribution when $s = 8$, $B = 1000$ and $\alpha = 0.01$.

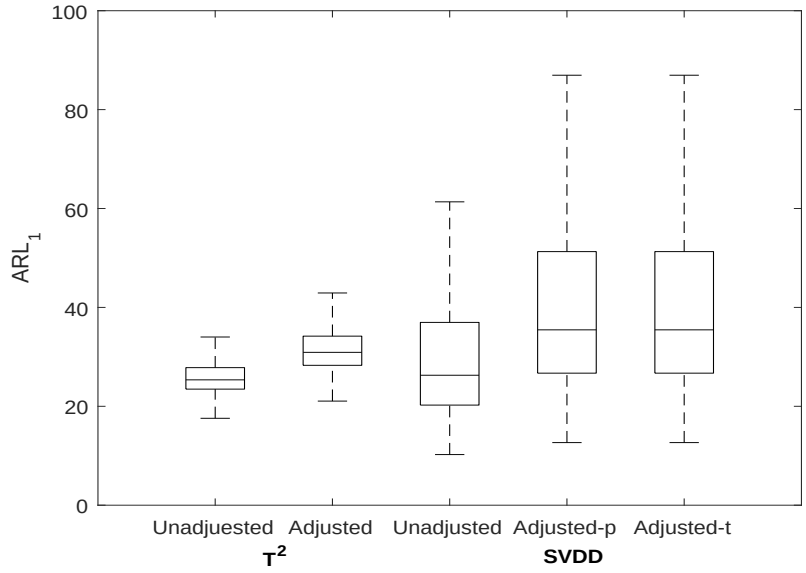


(a) $F_{\mathrm{BN}}(T, X|10, 1, 10, 2, 0.5)$

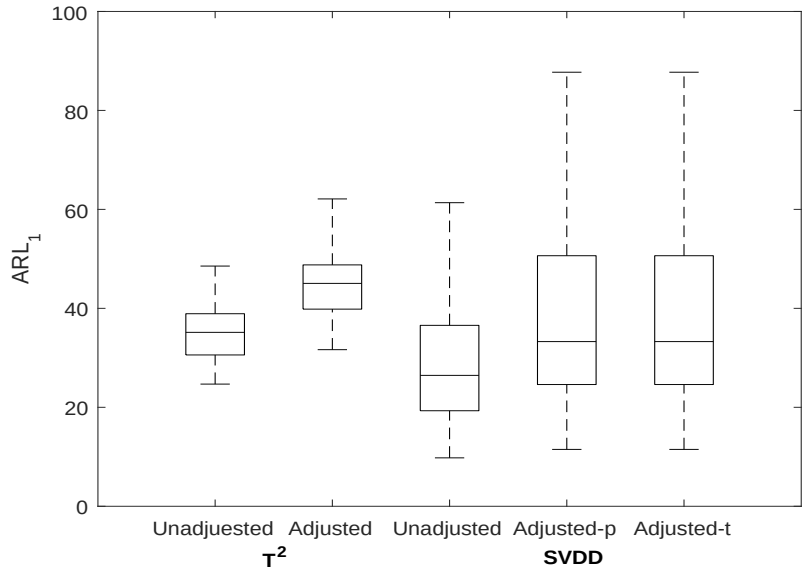


(b) $F_{\mathrm{SAT}}(T, X|4, 2.5, 100, 0.1, 0.5)$

Figure 6: Conditional in-control ARL_0 distribution of the Hotelling T^2 chart and the SVDD-based chart.

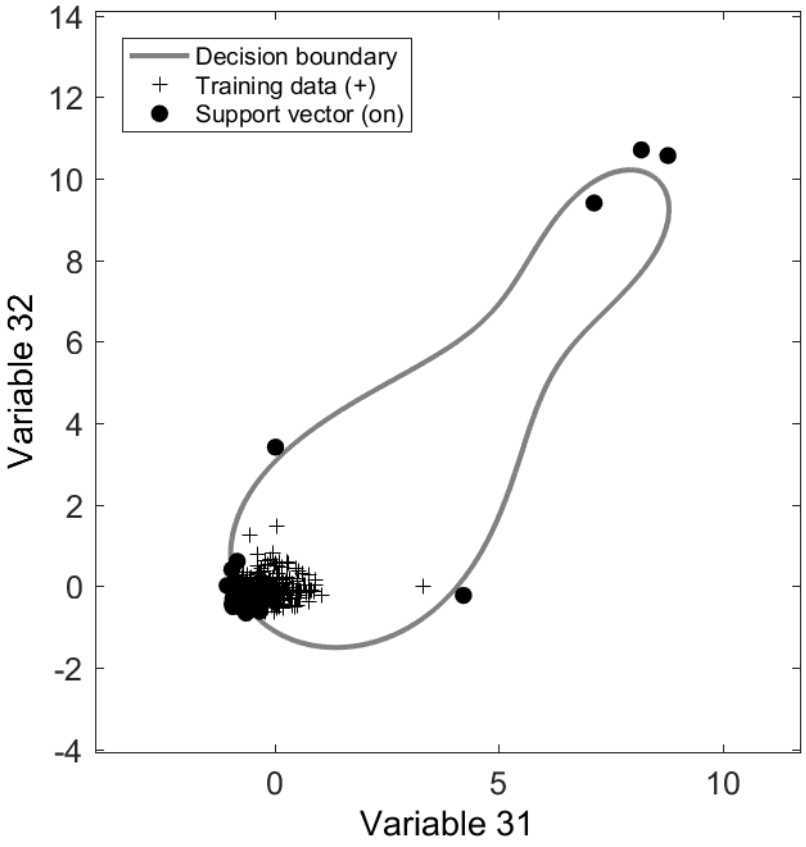


(a) $F_{\mathrm{BN}}(T, X|10, 1, 10, 2, 0.5)$

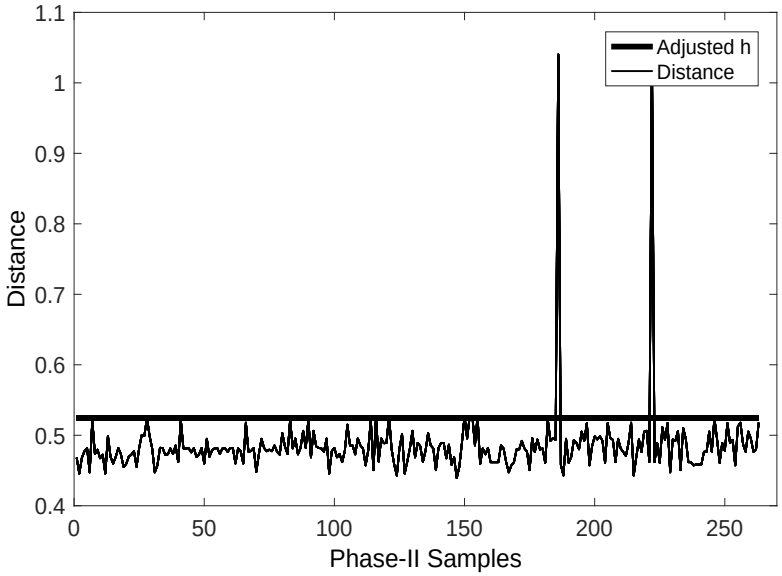


(b) $F_{\mathrm{SAT}}(T, X|4, 2.5, 100, 0.1, 0.5)$

Figure 7: Conditional out-of-control ARL_1 distribution of the Hotelling T^2 chart and the SVDD-based chart.



(a) The Phase-I sample



(b) The Phase-II sample

Figure 8: The SVDD-based chart applied to the morning of a semi-conductor manufacturing process