

Robust On-line Monitoring for Univariate Processes Based on Two Sample Goodness-of-fit Test

Chen Zhang, Nan Chen

Department of Industrial and Systems Engineering, National University of Singapore, Singapore
(zhangchen@nus.edu.sg & isecn@nus.edu.sg)

Abstract - On-line monitoring of quality variables and data streams raises attention in fields of quality management and statistical process control. Though much work has been dedicated on it in the literature, some challenges associated with designing distribution-free on-line control schemes are yet to be addressed well. This paper proposes a new distribution-free control chart for detecting both mean shifts and variance shifts in a univariate process based on nonparametric two sample goodness-of-fit test. It also integrates the self-starting feature by using data-dependent control limits set on-line instead of predetermined limits. Simulation study shows that the chart has satisfactory in-control run length performance given desired ARL and robust detection capability for general out-of-control changes, and is especially useful in short-run processes or start-up stages.

Keywords - Statistical process control (SPC), nonparametric, distribution-free, goodness-of-fit test

I. INTRODUCTION

In manufacturing, process quality is usually defined inversely proportional to process variability [1]. Statistical process control (SPC) schemes aim to monitor process variables to ensure that these variables do not shift away from the target values and the variance of them is acceptably small (usually smaller than certain threshold value which is called control limit). Usually, we can formulate SPC as the following change-point model. For an interested process variable X , we assume that there are m_0 reference samples X_{-m_0+1}, \dots, X_0 collected when the process is in control, and the i^{th} observation, X_i is collected over time following the change-point model:

$$X_i \stackrel{\text{i.i.d.}}{\sim} \begin{cases} F_0(x) & \text{for } i = -m_0 + 1, \dots, 0, 1, \dots, \tau, \\ F_1(x) & \text{for } i = \tau + 1, \dots \end{cases} \quad (1)$$

where τ is the unknown change point and $F_0(x)$ and $F_1(x)$ are unknown in-control (IC) and out-of-control (OC) distribution functions, respectively.

Since the first control chart was proposed by Shewhart in 1920s, a large amount of control charts, such like cumulative sum (CUSUM) charts and EWMA charts, have been developed to improve SPC performance. Readers could refer [1] and the references therein for a detailed review. However, most of current SPC schemes assume the process follows a normal distribution. This is simple but in practice the true process distribution is usually complex, such like being heavy tailed or skewed. Then charts with normal assumption will be inaccurate. To address this problem, distribution-free chart based on nonparametric

methods is useful. Its merit is that the practitioners need not assume a particular distribution or do in-control probability calculations for the underlying process and the associated performance remains valid for any distribution.

Current nonparametric charts are commonly based on “standardized ranks” of the observations, i.e., using empirical cumulative distribution function (ECDF) to represent the true process distribution. Reference [2] proposed a Shewhart-type chart based on ranks within rational groups. Reference [3] considered a CUSUM chart for group observations based on the Wilcoxon signed-rank statistic. Reference [4] proposed a CUSUM chart for individual observations based on sequential ranks. Though these CUSUM charts incorporate the sequential nature of SPC and are effective in detecting small, persistent process mean shifts, they need to know both IC and OC mean to set up the control limits. Reference [5] described an EWMA chart based on grouped signed-rank statistic. Reference [6] proposed an EWMA chart based on the goodness-of-fit test. Other relative work includes some change-point charts such like [7] and [8] based on the Mann-Whitney test and [9] based on the Kolmogorov-Smirnov and the Cramér-von Mises tests. For a more detailed discuss about the practical advantages of nonparametric charts and an overview of current univariate nonparametric charts, readers could refer to [10].

Most of these nonparametric charts focus on detecting process mean shifts. However, the change of process mean is usually masked or accompanied by an unsuspected change of the process variance. Monitoring more general process changes has become increasingly desirable. Among the aforementioned charts, [2] is designed for monitoring process variance when observations arrive in batches rather than as individuals. Reference [6] and [9] are designed for detecting both process mean and variance shifts. Unfortunately, they still have limitations and are inefficient in detecting decrease in process variance. Though decrease in variance receives little attention in the literature since in practice it usually indicates quality improvement instead of out of control, its influence on the chart performance should not be ignored when combined with mean shifts. For example, when the process undergoes a simultaneous mean shift and decrease in variance, the influence of the two change patterns on the chart statistics would counteract in the sense that the later delays the two charts above detecting the former. To avoid this problem, a chart designed to detect variance shifts in both

sides is preferable.

Furthermore, many existing nonparametric charts assume that the process IC parameters are known. This, however, is not realistic and these parameters need to be estimated based on reference samples. But collection of sufficient reference samples is sometimes infeasible or costs much, and practitioners usually want to monitor the process at start-up stages as early as possible. In these cases, the estimated parameters may be inaccurate and would add residual random variability to the run length distribution as a systematic bias, thereby compromising the chart performance [11]. One possible solution is to use self-starting chart, which allows monitoring processes from the start-up stages and not relying on many reference samples [12]. To be more specific, self-starting chart treats the reference samples as part of the ongoing data stream and examines the consistency through the whole process. Later [6], [8] and [9] successfully applied this idea to nonparametric charts. Recently, [13] proposed a framework of data-dependent self-starting monitoring where the control limit sequence is designed on-line based on current and previous observations rather than before monitoring. This idea ensures the chart has satisfactory IC run length especially for start-up stages and short-run processes.

This paper proposes a univariate distribution-free control chart to detect both process mean and variance shifts. Motivated by [14] and [15], a powerful two sample goodness-of-fit test is used to construct the monitoring statistic. To address the problem of insufficient reference sample set for distribution estimation, this chart integrates the self-starting feature by using data-dependent dynamic control limits [13]. Simulation study shows that given a desired ARL, this chart has satisfactory IC run length for any distribution, and robust detection power for general OC shifts.

II. METHODOLOGY

A. Two sample goodness-of-fit test

Let $S_0 = \{X_{-m_0+1}, \dots, X_0\}$ be *i.i.d* samples from an unknown distribution function $F_0(x)$, and $S_1 = \{X_1, \dots, X_n\}$ be *i.i.d* samples from an unknown distribution function $F_1(x)$. It is of interest to test whether S_0 and S_1 follow the same distribution, i.e., to test the hypothesis

$$\begin{aligned} H_0 : F_0(t) &= F_1(t) \quad \forall t \in (0, \infty). \\ H_1 : F_0(t) &\neq F_1(t) \quad \text{for some } t \in (-\infty, \infty). \end{aligned} \quad (2)$$

Without strong assumptions on $F_0(x)$ and $F_1(x)$, many nonparametric tests have been proposed, including the Kolmogorov-Smirnov test, the Anderson-Darling test, the Cramér-von Mises test, *etc.* Based on the nonparametric likelihood ratio (NLR), [14] proposed a new approach to construct a class of powerful one sample goodness-of-fit tests, which proved to be more powerful than many traditional tests. Later [6] used these tests in [14] to construct an EWMA control chart and showed the chart

was effective for detecting process mean shifts and increase in variance. However these one sample goodness-of-fit tests have limitations when applied into sequential SPC framework: If we consider the total sample set as $S = \{X_{-m_0+1}, \dots, X_n\}$, these tests only consider the distribution difference between S_1 and S but ignoring the difference between S_0 and S . To be more specific, if we treat S_0 as reference sample set and S_1 as sequential coming test sample set, the more samples S_1 accumulates, the more similarity S tends to have with S_1 , meaning that the less powerful the chart is to detect the difference between S_1 and $S(S_0)$. Consequently, unless the chart triggers an OC alarm soon after the process goes out of control, it tends to be decreasingly effective. In contrast, the two-sample goodness-of-fit test [15] overcomes this problem by considering the difference between S_1 and S , and between S_0 and S together. For each X_j , the two sample NLR test can be expressed as

$$L(X_j) = \sum_{i=0}^1 n_i \left(\hat{F}_i(X_j) \ln \frac{\hat{F}_i(X_j)}{\hat{F}(X_j)} + (1 - \hat{F}_i(X_j)) \ln \frac{1 - \hat{F}_i(X_j)}{1 - \hat{F}(X_j)} \right), \quad (3)$$

where n_i is the sample size of S_i . $\hat{F}_i(t)$ and $\hat{F}(t)$ are the ECDF's based on S_i and S respectively. Analogue to [15], by placing a weight $w_i(X_j)$ on $L(X_j)$ at different X_j and aggregating all the weighted information of $L(X_j)$, we could get a very powerful test statistic compared with many existing methods. Here we define the normalized geometric mean of $\hat{F}_i(X_j)$ and $1 - \hat{F}_i(X_j)$ as the weight function, i.e., $w_i(X_j) = [\hat{F}_i(X_j)(1 - \hat{F}_i(X_j))]^{-0.5}$, and summarize $L(X_j)$ for every X_j as

$$Z = \sum_{j=1}^n \sum_{i=0}^1 n_i \left\{ \sqrt{\frac{\hat{F}_i(X_j)}{1 - \hat{F}_i(X_j)}} \ln \frac{\hat{F}_i(X_j)}{\hat{F}(X_j)} + \sqrt{\frac{1 - \hat{F}_i(X_j)}{\hat{F}_i(X_j)}} \ln \frac{1 - \hat{F}_i(X_j)}{1 - \hat{F}(X_j)} \right\}. \quad (4)$$

Note that other weight functions could also be considered, but how to choose the weight is not the focus of this paper.

B. A Distribution-free Control Chart

Now we deploy the two sample goodness-of-fit test into on-line sequential monitoring. When a new observation X_n comes, a sliding window is used to update S_1 , i.e., $S_1^n = \{X_{n-w+1}, X_n\}$, the rest samples are used as updated $S_0^n = \{X_{-m_0+1}, \dots, X_{n-w}\}$ in the self-starting view. Then the fix-sample test in (4) could be used to get Z_n . Figure 1 illustrates the framework of the chart. To address more on recent samples, similar to [6], we consider the weighted empirical distribution for S_1^n and replace $\hat{F}_1(X_j)$ as

$$\hat{F}_1^\lambda(X_j) = \frac{1}{n_1} \sum_{k=n-w+1}^n (1 - \lambda)^{n-k} \mathbf{I}(X_k \leq X_j)$$

with

$$n_1 = \sum_{k=n-w+1}^n (1-\lambda)^{n-k},$$

where λ could be regarded as a weighting parameter commonly used in EWMA charts. Then we get the monitoring statistic Z_n^λ as

$$Z_n^\lambda = \sum_{j=1}^n \frac{(1-\lambda)^{n-j}}{a_{\lambda,n}} \sum_{i=0}^1 n_i \left(\sqrt{\frac{\hat{F}_i^*(X_j)}{1-\hat{F}_i^*(X_j)}} \ln \frac{\hat{F}_i^*(X_i)}{\hat{F}_i^*(X_j)} + \sqrt{\frac{1-\hat{F}_i^*(X_j)}{\hat{F}_i^*(X_j)}} \ln \frac{1-\hat{F}_i^*(X_j)}{1-\hat{F}_i^*(X_j)} \right) \quad (5)$$

with $\hat{F}_1^*(X_j) = \hat{F}_1^\lambda(X_j)$ and $\hat{F}_0^*(X_j) = \hat{F}_0(X_j)$. Note that Z_n^λ is nonparametric since it only uses the rank information of X_j 's rather than the magnitude information. Therefore the chart based on two sample goodness-of-fit test is distribution-free in the sense that its IC run-length distribution is the same for all process distributions.

Based on this point, we could determine a certain control limit such that the chart has satisfactory run-length performance regardless of process distribution. Usually a chart is considered to be satisfactory if its IC run length distribution is close to a geometric distribution [16]. In addition, it might be unacceptable if the specified IC ARL (denoted as ARL_0) is attained with an elevated probability of false alarms in short runs as compared to the geometric distribution. Here inheriting the idea of [13], we aim to find the control limit sequence $\{H_n\}$ such that the conditional false alarm probability at each point given that there is no false alarm before is a pre-specified constant α , i.e., for $1 < i < n$ and $n > 1$,

$$\begin{aligned} P(Z_1^\lambda > H_1(\alpha)) &= \alpha, \\ P(Z_n^\lambda > H_n(\alpha) | Z_i^\lambda < H_i(\alpha)) &= \alpha. \end{aligned} \quad (6)$$

This is equivalent to perform a hypothesis test with the type-I error α at each time point n . Based on $H_i(\alpha)$, we define the following charting procedure, termed as distribution-free two sample goodness-of-fit chart (abbreviated as 2SGoF), with the run-length

$$RL = \min\{n; Z_n^\lambda \geq H_n(\alpha), n \geq 1\}. \quad (7)$$

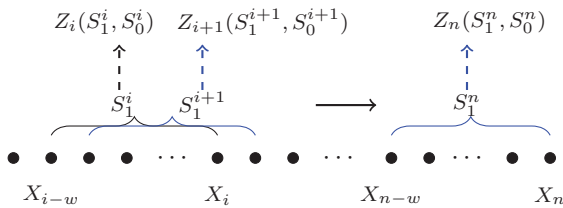


Fig. 1. The framework of sliding window method for on-line monitoring. Each circle represents one sample, w is the sliding window length.

III. PERFORMANCE STUDY

In this section, we present some simulation results about the performance of 2SGoF and compare it with some related approaches mentioned in the literature.

We consider the following distributions in our study: (1) normal distribution; (2) t_5 , the student t distribution with 5 degrees of freedom; (3) χ_3^2 , the chi-square distribution with 3 degrees of freedom. Without loss of generality, for each distribution, we set the IC mean $\mu_0 = 0$, variance $\sigma_0^2 = 1$, reference sample size $m_0 = 100$, sliding window $w = 58$, and the desired $ARL_0=200$. For OC settings, we consider the following scenarios: (1) the process mean shift of size δ , i.e., $\mu_1 = \mu_0 + \delta$; (2) the process variance shift of magnitude λ^2 , i.e., $\sigma_1^2 = \lambda^2 \sigma_0^2$.

Other process change scenarios are also considered and show that the general conclusions given below do not change. These additional simulation results are available from the authors on request. All the results in this section are obtained from 10,000 replications.

A. In-control Performance

First, we study the IC run length distribution of the 2SGoF chart in terms of ARL_0 , standard deviation of the run length (SDRL) and the false-alarm rate during the first 30 observations, i.e., $FAR = P(RL \leq 30)$. Two values of λ , 0.05 and 0.1 are considered for a better illustration. In our settings with $ARL_0 = 200$, if the run length distribution is geometric, α should be 0.005 with the corresponding $SDRL = ARL_0$ and $FAR = 0.140$. As Table I shows, for these three distributions, 2SGoF always has satisfactory IC performance with ARL_0 values almost 200, and its SDRL and FAR values are close to the ones with geometric distribution. This can be understood as the IC distribution of 2SGoF is exactly geometric, thereby guaranteeing its feasibility for distribution-free process monitoring.

B. Out-Control Performance

Next, we analyze the OC performance of 2SGoF and compare it with some other self-starting work for monitoring both process mean and variance, including the nonparametric likelihood ratio-based EWMA chart (NLE) [6], the parametric likelihood ratio Change Point model (ChangePt) [17], and the nonparametric Cramér-von Mises (CvM) model [9]. $m_0 = 100$ satisfies the requirements for

TABLE I
IC PERFORMANCE OF THE 2SGoF WITH $N_{0,1}$, t_5 AND χ_3^2
OBSERVATIONS WHEN $m_0 = 100$

	$\lambda = 0.1$			$\lambda = 0.05$		
	ARL_0	SDRL	FAR	ARL_0	SDRL	FAR
geometric	200.0	200.0	0.140	200.0	200.0	0.140
$N_{0,1}$	200.0	199.4	0.141	199.0	196.2	0.142
t_5	198.4	196.1	0.142	198.0	195.8	0.144
χ_3^2	197.6	195.4	0.146	195.8	193.4	0.151

TABLE II

OC ARL COMPARISON IN DETECTING LOCATION SHIFTS WHEN $m_0 = 100$ AND $\lambda = 0.05$; NUMBERS IN PARENTHESES ARE SDRL VALUES.

	δ	2SMGoF	NLE	ChangePt	CvM
$N_{0,1}$	0.25	92.3(114.9)	118.6(153.1)	138.6(153.4)	117.5(128.8)
	0.5	29.1(29.3)	41.7(54.6)	51.6(54.4)	35.3(38.5)
	1.0	11.1(5.5)	14.9(6.93)	13.3(7.9)	10.3(9.5)
	2.0	5.3(1.94)	6.3(2.8)	4.7(1.8)	4.3(6.1)
	4.0	3.6(1.0)	3.4(5.6)	2.1(0.6)	2.9(5.1)
t_5	0.25	116.8(138.3)	121.5(145.5)	170.4(174)	122.2(133.9)
	0.5	37.7(39.6)	57.6(74.3)	93.2(99.8)	44.2(50.9)
	1.0	13.9(7.6)	19.5(10.2)	24.2(17.1)	12.3(11.5)
	2.0	6.6(2.6)	9.1(3.9)	7.6 (3.4)	4.9(6.6)
	4.0	4.1(1.2)	4.2(1.8)	3.0(1.0)	3.1(5.7)
χ_3^2	0.25	54.1(75.4)	76.4(115.2)	197.4(147.1)	99.4(117.2)
	0.5	18.7(9.2)	28.9(12.8)	109.7(109.2)	23.7(20.7)
	1.0	10.0(3.9)	17.2(6.4)	22.7(11.7)	7.9(7.7)
	2.0	5.9(2.0)	9.1(3.5)	6.5(2.4)	4.1(6.0)
	4.0	3.9(1.2)	4.2(1.7)	2.5(0.7)	3.0(5.6)

starting other charts. We fix $\tau = 25$ and only consider the steady-state ARL, meaning that any series where an OC signal occurs before the true change point τ is discarded. To give a fair comparison, we also set the $\lambda = 0.05$ for NLE chart. As to ChangePt, since it is designed with normal distribution assumption, for t_5 and χ_3^2 we artificially tune its control limits to ensure that its ARL_0 is still equal to the normal one, i.e., $ARL_0 = 200$. This kind of adjustments could only be used for simulation comparison, but not applicable in practice since usually the process is unpredictable.

Table II illustrates the simulation results for the charts under a mean shift with $\delta = 0.25, 0.5, 1, 2, 4$. For small shifts ($0 \leq \delta < 1$), 2SGoF chart outperforms the other three charts dominantly by a large margin. For moderate and large mean shift ($\delta = 1, 2, 4$), though 2SGoF responds slightly slower than CvM (with at most 1 time point delay), it still performs generally better than the other two charts.

Table III compares the charts in detecting either an increasing or decreasing variance shift. We observe that 2SGoF is close to the best in general and has robust properties. When variance increases, NLE always has the fastest response among the four charts, and 2SGoF performs better than CvM and ChangePt. When variance decreases, their performance is opposite: the three nonparametric charts have less detection power than ChangPt and are specially inefficient for $\lambda^2 = 0.75$. But 2SGoF performs alternatively better than NLE and CvM.

IV. CONCLUSION

This paper presents a univariate distribution-free control chart based on the two sample likelihood ratio test. As analyzed in [14] and [15], this likelihood ratio test has better performance than traditional nonparametric goodness-of-fit test statistics. The proposed method has superior

TABLE III

OC ARL COMPARISON IN DETECTING VARIANCE SHIFTS WHEN $m_0 = 100$ AND $\lambda = 0.05$; NUMBERS IN PARENTHESES ARE SDRL VALUES.

	λ^2	2SMGoF	NLE	ChangePt	CvM
$N_{0,1}$	0.25	33.4(25.1)	50.2(68.8)	17.2(8.9)	104.9(97)
	0.5	154.9(176.7)	299.1(346.3)	57.6(60.2)	246.8(218.7)
	0.75	225.6(196.2)	332(292.9)	157.1(166.1)	207.9(196.7)
	1.25	140.91(144.9)	111.5(156.3)	184.8(185.1)	131.9(121.5)
	1.5	99.1(113.5)	63.4(142.7)	143.1(155.5)	107.7(128.1)
	2.0	49.5(59.7)	29.4(27.6)	57.2(64.5)	76.3(103.2)
	4.0	15.4(9.6)	11.2(10.1)	11.9(8.6)	29.2(39.6)
t_5	0.25	41.3(41.9)	71.2(109.6)	19.5(12.9)	153.6(127.9)
	0.5	171.2(183.5)	265.6(280.2)	64.8(76.9)	274.6(215.7)
	0.75	214.8(194.2)	267.1(245)	158.9(165.8)	240.5(201.7)
	1.25	159.4(161.5)	136.6(153.6)	185.6(184.9)	158.5(157.5)
	1.5	123.7(134.7)	92.4(115)	146.4(161.8)	134.4(146.8)
	2.0	72.8(85.2)	47.1(57.4)	80.1(103.4)	95.8(106.7)
	4.0	21.5(17.3)	15.9(10.5)	17.6(17.5)	42.7(46.4)
χ_3^2	0.25	21.6(8.1)	35.8(12.1)	26.1(17.6)	46.9(31.7)
	0.5	53.1(73.4)	66.6(89.3)	84.2(100.2)	185.9(168)
	0.75	179.1(187.4)	325.9(339.6)	169.3(173.3)	244.3(204.3)
	1.25	89.31(106.7)	67.7(45.9)	182.5(180.9)	130.7(133.7)
	1.5	43.1(50.5)	22.7(18.8)	165.4(172.8)	95.7(107.5)
	2.0	20.6(15.9)	9.1(3.5)	105.3(127.6)	48.4(58.3)
	4.0	10.2(5.5)	7.4(4.1)	24.7(25.3)	20(21.1)

performance in detecting the changes in the mean and variance in both increasing and decreasing directions, regardless of the IC distributions. Furthermore, we also integrate self-starting feature into the chart by using dynamic control limits. The self-starting method can ensure that the proposed method has desired IC performance without excessive false alarms even with small reference sample size. We compare its performance with several nonparametric charts discussed in the literature. The results show that our method is very promising, especially in detecting small mean shifts.

There are also some possible extensions to our current method. For example, how to generalize our proposed method to multivariate process control with efficient detecting capability worth further study. It is also interesting to study how the weighting function influences the performance of the charts, and how to identify the optimal weighting function.

REFERENCES

- [1] D. C. Montgomery, *Introduction to statistical quality control*. Wiley New York, 1991.
- [2] L. Jones-Farmer, C. W. Champ *et al.*, "A distribution-free phase i control chart for subgroup scale," *Journal of quality technology*, vol. 42, no. 4, pp. 373–387, 2010.
- [3] S. T. Bakir and M. R. Reynolds, "A nonparametric procedure for process control based on within-group ranking," *Technometrics*, vol. 21, no. 2, pp. 175–183, 1979.
- [4] D. McDonald, "A cusum procedure based on sequential ranks," *Naval Research Logistics (NRL)*, vol. 37, no. 5, pp. 627–646, 1990.

- [5] R. W. Amin and A. J. Searcy, "A nonparametric exponentially weighted moving average control scheme," *Communications in Statistics-Simulation and Computation*, vol. 20, no. 4, pp. 1049–1072, 1991.
- [6] C. Zou and F. Tsung, "Likelihood ratio-based distribution-free ewma control charts," *Journal of Quality Technology*, vol. 42, no. 2, p. 174, 2010.
- [7] S. Chakraborti, M. A. van de Wiel *et al.*, *A nonparametric control chart based on the Mann-Whitney statistic*. Institute of Mathematical Statistics, 2008.
- [8] C. Zhou, C. Zou, Y. Zhang, and Z. Wang, "Nonparametric control chart based on change-point model," *Statistical Papers*, vol. 50, no. 1, pp. 13–28, 2009.
- [9] G. J. Ross and N. M. Adams, "Two nonparametric control charts for detecting arbitrary distribution changes." *Journal of Quality Technology*, vol. 44, no. 2, 2012.
- [10] S. Chakraborti, P. Van Der Laan, and S. T. Bakir, "Nonparametric control charts: An overview and some results," *Journal of Quality Technology*, vol. 33, no. 3, pp. 304–315, 2001.
- [11] L. A. Jones, C. W. Champ, and S. E. Rigdon, "The performance of exponentially weighted moving average charts with estimated parameters," *Technometrics*, vol. 43, no. 2, 2001.
- [12] D. M. Hawkins, P. Qiu, and C. W. Kang, "The changepoint model for statistical process control," *Journal of Quality Technology*, vol. 35, no. 4, pp. 355–366, 2003.
- [13] N. Chen, X. Zi, and C. Zou, "A distribution-free multivariate control chart," *under revision*, 2013.
- [14] J. Zhang, "Powerful goodness-of-fit tests based on the likelihood ratio," *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, vol. 64, no. 2, pp. 281–294, 2002.
- [15] —, "Powerful two-sample tests based on the likelihood ratio," *Technometrics*, vol. 48, no. 1, pp. 95–103, 2006.
- [16] D. M. Hawkins and D. H. Olwell, *Cumulative sum charts and charting for quality improvement*. Springer, 1998.
- [17] D. M. Hawkins and K. Zamba, "Statistical process control for shifts in mean or variance using a changepoint formulation," *Technometrics*, vol. 47, no. 2, 2005.