# A Nonparametric Change-Point Control Chart

DOUGLAS M. HAWKINS and QIQI DENG

*School of Statistics, University of Minnesota, Minneapolis, MN 55455-0493*

The assumption of fully known in-control distributions has long been recognized as an idealization, at best approximately true. Recent development of normal-based change-point methods has allowed the assumption of exactly known in-control mean and variance to be relaxed, but retained the assumption of normality. In this paper, we develop a nonparametric tool based on the change-point model for statistical process control. This method is shown to perform well, even beating the parametric approach for small to moderate shifts in normal data, and to involve relatively light computation.

Key Words: Chart Performance; Cumulative Methods; Rank Sum Test.

In statistical process control (SPC), one major concern is whether there has been a change of the distribution from the target in the process. To answer this question, we have to distinguish the variation due to real change of distribution (assignable causes) from that due to random error (chance causes). Many kinds of control charts serve this purpose. When there is no change and the process is in control, the probability of a false signal from the chart should be controlled at a low and known false-alarm rate. And when there is some change and the process is out of control, a good chart should detect it as soon as possible. The performance of a chart is usually measured by average run length (ARL)—the expected number of samples or subgroups to be collected before the first signal.

Control charts often assume that data come from some parametric distribution, most commonly the normal distribution. When the underlying process is unknown or known not to be normal, these charts may not be appropriate. Woodall and Montgomery (1999) and Woodall (2000) have provided some discussion on this issue. Along with other literature, they motivate the development and use of nonparametric control charts. Chakraborti et al (2001) gave a comprehensive review of nonparametric control charts. Amin et al. (1995) developed Shewhart-type control charts assuming the location parameter (mean or median) is known. Recently Jones-Farmer et al. (2009) proposed a Shewhart-type Phase I method based on ranks within rational groups. Some other Shewhart charts, for example, Willemain and Runger (1996) and Janacek and Meikle (1997) are set up on some attribute of the in-control distribution, like the mean, median, percentile, or interquartile range estimated from a reference sample. Bakir and Reynolds (1979) considered a CUSUM chart based on the Wilcoxon signed-rank (WSR) statistic. McDonald (1990) proposed a CUSUM based on "sequential ranks". These CUSUM methods incorporate the sequential nature of SPC and are thus effective in detecting small, persistent change. However, they require knowledge of both the in-control and out-of-control location parameters to set up the reference value. An EWMA-type chart was provided by Amin and Search (1991), who described a GSR-EWMA chart based on a grouped signed-rank statistic (GSR). This chart too requires known in-control parameters to set up its control limits. Another EWMA-based yet nonparametric method has been proposed by Zou and Tsung (2010). Hackel and Ledolter (1991) used "standardized ranks" of the observations relative to an in-control distribution. When the in-control distribution is unknown, they suggested using the ranking on reference data to estimate it, which may lead to a substantially larger in-control ARL than nominal.

Dr. Hawkins is a Professor at University of Minnesota. He is an ASQ Fellow. His email address is dhawkins@umn.edu.

Dr. Deng is a Senior Biostatistician with Boehringer Ingelheim Pharmaceuticals. Her email address is qiqi.deng@boehringer-ingelheim.com.

All these methods require prior knowledge of the in-control parameter or at least a large reference sample to get an adequate estimate of the parameter(s). This reference sample is gathered during a separate Phase I study in which the process is believed to be operating in control, but otherwise there is no special effort to get unrealistically tight readings (Montgomery (2005), section 4.1.7). Jones et al. (2001, 2004) and Jensen et al. (2006) discussed the effect of estimating the in-control behavior of process readings on the performance of control charts and pointed out that, with a moderately sized reference sample, control charts based on the estimated in-control behavior will behave quite differently from desired. Specifically, estimation usually substantially increases the in-control ARL from the nominal value, and also reduces the sensitivity to change when the process goes out of control.

Recently, Hawkins et al. (2003) (henceforth, HQK) proposed a change-point control chart for normally distributed data. Instead of comparing the subsequent observations with a known or estimated target value, their method treats the reference samples as part of the ongoing data stream and examines the consistency through the whole process. Lai (2001) listed a variety of change-point models for sequential detection in settings where at least some in-control parameters were assumed to be known in advance. The HQK formulation however differs in not requiring advance knowledge about any parameter. The HQK method not only avoids the tenuous "known-parameter" assumption but always maintains the desired in-control ARL.

Zhou et al (2009) developed a nonparametric change-point model motivated by HQK, but did not follow the same methodology as HQK. Their scheme and its performance will be discussed below. In this paper, we will develop a direct nonparametric parallel of HQK with modest computational needs.

## A Review of a Nonparametric Fixed-Sample-Size Approach

SPC tools are conventionally used in one of two settings. In Phase I, we have a data set of fixed size, of which the primary purpose is to estimate the in-control properties of the process readings. Phase II involves a steady stream of incoming readings but conventionally does not involve any further refinement of the estimates of in-control process behavior. To develop our nonparametric Phase II methodology, it is helpful to first sketch the change-point formulation where you have a static data set.

Assume $X_1, X_2, \ldots, X_\tau, X_{\tau+1}, \ldots, X_n$ are independent, continuous random variables with statistical distribution

$$X_i \sim F(x), \qquad \text{for } i = 1, 2, \ldots, \tau; \tag{1}$$

$$X_i \sim F(x-\theta), \qquad \text{for } i = \tau+1, \ldots, n. \tag{2}$$

The parameter $\theta$ represents a shift in location occurring after the change point $\tau$. Both $\theta$ and $\tau$ are assumed unknown. Testing whether the process has shifted corresponds to the hypothesis test

$$H_0 : \theta = 0, \quad H_a : \theta \neq 0$$

or, equivalently, that the change point $\tau$ lies outside the range $1 \ldots n$.

Pettitt (1979) proposed a U-statistic based on the Mann–Whitney two-sample test, Let

$$D_{ij} = \operatorname{sgn}(X_i - X_j) = \begin{cases} 1 & \text{if } X_i > X_j \\ 0 & \text{if } X_i = X_j \\ -1 & \text{if } X_i < X_j. \end{cases}$$

The $U$-statistic $U_{k,n}$ is defined as an antisymmetric function of the $D_{ij}$,

$$U_{k,n} = \sum_{i=1}^{k} \sum_{j=k+1}^{n} D_{ij}, \qquad 1 \le k \le n-1. \tag{3}$$

A well-known result (Conover (1999)) connects $U_{k,n}$ to the rank sum statistic: Let $R_i$ be the rank of $X_i$ within $X_1, X_2, \ldots, X_n$. From the two equivalent formulas of the Mann–Whitney statistic, it is known that

$$U_{k,n} = 2\sum_{i=1}^{k} R_i - k(n+1).$$

It follows that

$$E(U_{k,n}) = 0, \quad \operatorname{Var}(U_{k,n}) = \frac{k(n-k)(n+1)}{3}. \tag{4}$$

Note that the variance of $U_{k,n}$ depends on the putative split point $k$, being a maximum when $k = n/2$. Schechtman and Wolfe (1981, 1984) proposed standardizing the $U_{k,n}$ to constant variance, defining

$$T_{k,n} = \frac{U_{k,n}}{\sqrt{k(n-k)(n+1)/3}}. \tag{5}$$

From the asymptotic normality of Mann–Whitney statistics, $T_{k,n} \sim N(0,1)$ when $k$ and $n-k$ both go to $\infty$.

A test statistic for the presence of a change point, and an estimate of its time of occurrence, are then given by maximizing this statistic over $k$,

$$T_{\max,n} = \max_{1\le k\le n-1} |T_{kn}| \tag{6}$$
$$\hat{\tau}_T = \operatorname*{argmax}_{1\le k\le n-1} |T_{k,n}|.$$

It is helpful at this point to look at the standard parametric change-point statistic—see, for example, Hawkins (1977). Suppose the distribution $F$ is the normal distribution with mean $\mu$ and known variance $\sigma^2$ and define

$$\bar{X}_{i,j} = \sum_{k=i+1}^{j} X_k/(j-i)$$
$$Y_{k,n} = \frac{\bar{X}_{0,k} - \bar{X}_{k,n}}{\sigma(k^{-1} + (n-k)^{-1})}.$$

In the null case, these $Y_{k,n}$ are N(0, 1) and turn out to have the same correlation structure as the $T_{k,n}$ (Deng (2009)). Thus, except for small $k$ or $n-k$, the parametric $Y_{k,n}$ statistics have effectively the same in-control statistical properties as their nonparametric counterparts $T_{k,n}$.

The parametric test statistic for a change is

$$Y_{\max,n} = \max_{1\le k\le n-1} |Y_{kn}|. \tag{7}$$

Csörgő, and Horváth (1997) discussed the large-sample asymptotics of $Y_{\max,n}$ and $T_{\max,n}$, showing that both tend to infinity with probability 1, but slowly, and that they have the same asymptotic extreme-value distribution.

In the parametric case, if $\sigma$ is unknown, it is replaced by $\hat{\sigma}_{k,n}$, the pooled within-segment standard deviation of the two putative segments, replacing $Y_{k,n}$ by the two-sample $t$ test for equality of mean of the two "samples" formed by the split after observation $k$.

$$Z_{k,n} = \frac{\bar{X}_{0,k} - \bar{X}_{k,n}}{\hat{\sigma}_{k,n}(k^{-1} + (n-k)^{-1})}$$
$$Z_{\max,n} = \max_{1\le k\le n-1} |Z_{kn}|.$$

The strong consistency of $\hat{\sigma}$ implies that this studentized version $Z_{\max,n}$ has the same asymptotic behavior as the known-$\sigma$ form $Y_{\max,n}$.

## A Nonparametric Tool for Phase II SPC

Turning now to Phase II SPC, the data set is no longer fixed, but grows as long as the process is believed to still be in control. Paralleling the HQK methodology, as each new observation accrues, for our nonparametric change-point control method, we look at every previous time as a potential change point, carrying out a two-sample Mann–Whitney test between the observations preceding, and those following that time. If the maximum of these statistics exceeds a specified control limit $h_n$, the method signals that a change has occurred and estimates the time of change as the maximizing split point. In symbols, defining a sequence of control limits $h_n$, for each $n$, we

- Compute $T_{\max,n}$, the maximized split statistic.
- If $T_{\max,n} \le h_n$, then conclude that the process is in control and continue to the next reading.
- If $T_{\max,n} > h_n$, then conclude that the process has shifted and stop the process for diagnosis. Estimate the epoch of the shift by the maximizing $k$.

This raises the issue of the control-limit sequence $h_n$. As in the parametric HQK setting, we define these by fixing the conditional probability of a false alarm at any observation, given that there was no false alarm at all the previous test. The ideal $h_n$ should be set such that this conditional probability of a false alarm at any $n$ was a constant $\alpha$. If this is done, then the run length follows a geometric distribution with probability $\alpha$, making the in-control average-run length $1/\alpha$ from any in-control starting point. We refine the notation to $h_{n,\alpha}$ to reflect this dependence of the sequence on the false-alarm probability.

The $h_{n,\alpha}$ are defined by the joint null distribution of the $T_{k,n}$. However, the exact distribution is discrete and varies as $n$ or $k$ changes and requires finding out all the possible arrangements of the ranks. That would be a nightmare even for moderate $n$ and unthinkable for large $n$, as discussed in Schechtman (1982), and this problem is even more severe in Phase II study. We will attack it by large-scale simulation.

In principle, it is possible to start testing for a change from the third observation because the change-point model does not rely on parameter estimates. However, because of the discrete nature of the distribution of $T_{\max,n}$, with short sequences, it is impossible to have control limits with the small $\alpha$ conventionally used in SPC settings, and so to get to these low false-alarm rates, we need to have some minimum number of "warm-up" cases before we start

TABLE 1. Cutoffs $h_{n,\alpha}$ for Sample Size $n$ and Varying IC ARL Starting at Sample 15

| $n$ | In-Control ARL 50 | 100 | 200 | 500 | 1000 | 2000 |
|---|---|---|---|---|---|---|
| 15 | 2.700 | 2.848 | 2.947 | 3.069 | 3.181 | 3.229 |
| 16 | 2.615 | 2.767 | 2.910 | 3.047 | 3.142 | 3.244 |
| 17 | 2.535 | 2.718 | 2.862 | 3.043 | 3.163 | 3.247 |
| 18 | 2.535 | 2.694 | 2.860 | 3.034 | 3.183 | 3.277 |
| 19 | 2.500 | 2.695 | 2.869 | 3.054 | 3.186 | 3.296 |
| 20 | 2.488 | 2.699 | 2.851 | 3.059 | 3.203 | 3.311 |
| 22 | 2.468 | 2.692 | 2.862 | 3.082 | 3.228 | 3.355 |
| 24 | 2.469 | 2.676 | 2.870 | 3.096 | 3.249 | 3.389 |
| 26 | 2.452 | 2.686 | 2.875 | 3.108 | 3.269 | 3.415 |
| 28 | 2.455 | 2.686 | 2.883 | 3.121 | 3.283 | 3.437 |
| 30 | 2.453 | 2.684 | 2.879 | 3.130 | 3.297 | 3.453 |
| 35 | 2.452 | 2.687 | 2.894 | 3.149 | 3.324 | 3.487 |
| 40 | 2.447 | 2.689 | 2.900 | 3.162 | 3.342 | 3.511 |
| 45 | 2.453 | 2.690 | 2.906 | 3.171 | 3.356 | 3.529 |
| 50 | 2.451 | 2.691 | 2.908 | 3.178 | 3.365 | 3.542 |
| 60 | 2.452 | 2.694 | 2.914 | 3.188 | 3.379 | 3.560 |
| 70 | 2.452 | 2.694 | 2.917 | 3.194 | 3.388 | 3.570 |
| 80 | 2.453 | 2.696 | 2.918 | 3.199 | 3.394 | 3.579 |
| 90 | 2.452 | 2.696 | 2.920 | 3.200 | 3.399 | 3.584 |
| 100 | 2.453 | 2.697 | 2.922 | 3.203 | 3.402 | 3.591 |
| 125 | | 2.698 | 2.923 | 3.206 | 3.409 | 3.599 |
| 150 | | 2.697 | 2.924 | 3.209 | 3.411 | 3.603 |
| 200 | | 2.699 | 2.926 | 3.210 | 3.415 | 3.610 |
| 250 | | 2.700 | 2.927 | 3.212 | 3.416 | 3.610 |
| 300 | | 2.704 | 2.926 | 3.215 | 3.420 | 3.616 |
| 500 | | | 2.927 | 3.213 | 3.417 | 3.612 |
| 1000 | | | 2.927 | 3.214 | 3.418 | 3.612 |

the testing. Finding the exact probability distribution of $T_{k,n}$ for various small values of $n$ shows that a warm-up of 14 cases is the smallest value suitable for routine SPC use. It allows $\alpha$ values as small as 1/3432, which is small enough for SPC use. With 14 warm-up data, the method starts actual monitoring from the 15th process reading. This is the method we have implemented.

Forty million sequences of length 1000 were simulated to find control limits up to $n = 1000$. These control limits are presented in Table 1, the columns corresponding to in-control average run lengths of 50, 100, 200, 500, 1000, and 2000, respectively. The table suggests that the cutpoints stabilize to constant values for increasing $n$. Not all $n$ values are listed explicitly; missing entries can be found by interpolation

TABLE 2. Comparing the Parametric (P) and Nonparametric (NP) Cutpoints

| | $\alpha$ 0.005 | 0.002 | 0.001 |
|---|---|---|---|
| P | 3.248 | 3.570 | 3.794 |
| NP | 2.926 | 3.250 | 3.455 |

or by carrying entries forward. For example, there is an entry for $n = 60$ and one for $n = 70$; little error is introduced by using the $n = 60$ entry for all values between 60 and 70, and some slight improvement in accuracy can be gained by interpolating.

In view of the close similarity between the asymptotics of the parametric and nonparametric procedures in the situation of a fixed-size data set, it is interesting to ask whether their asymptotic control limits in the Phase II setting are close. A comparison of the control limits for $n = 200$ given in Table 2 shows that they are markedly different. The control limits for the nonparametric procedure correspond to parametric procedures with more than double the false-alarm rate. That two approaches with such similar Phase I asymptotics should differ so markedly in Phase II is surprising. The difference must be traced to the impact of very short segments, as this is the only substantive difference in the large-sample in-control distributions of the underlying statistics.

With this material in place, we can look more closely at the nonparametric chart proposed by Zhou et al (2009). This method starts with the $T_{k,n}$ statistics defined earlier. But rather than test these statistics directly, they are replaced by smoothed versions generated by an exponentially weighted moving average, defined by

$$E_{k,n} = (1-\lambda)E_{k,n-1} + \lambda T_{k,n}$$
$$W_{\max,n} = \max_{1\le k\le n-1} |E_{kn}|,$$

a change point being declared if $W_{\max,n}$ exceeds a control limit chosen to fix the conditional probability of a false alarm at a constant $\alpha$.

The constant $\lambda$, as usual with an EWMA, controls the degree of smoothing. Zhou et al. (2009) suggest the value $\lambda = 0.2$ as a good general-purpose choice.

## Computational Issues

The parametric HQK algorithm requires storage of a table of running sums and sums of squared deviations—$2n$ words in all. Computationally efficient numerically stable updates of these summary tables are fast, and as $n$ increases, the only significant computation is the search across previous values for the location of the maximizing change point. An instruction count shows that this operation uses $12n$ arithmetic operations and so scales linearly.

Turning to the nonparametric case, at first glance, the computational issue seems far more intractable. Calculating the rank sum requires ordering the data. This is an $O(n \log n)$ operation (in other words, the amount of computing required to do it increases with $n$ at a rate of $n \log n$), while the definitional form of $U_{k,n}$ is $O(n^2)$, so its computation goes up as the square of the size of the data set. In the ongoing Phase II analysis with new data coming in continuously, all orderings change as each new observation is added and the rankings need continual updating. A naive implementation of the definitional form of $T_{\max,n}$ would therefore lead to a computational explosion with large $n$.

More careful analysis, however, shows that the Mann–Whitney statistic has a simple update: for $1 \le k \le n$,

$$\begin{aligned} U_{k,n+1} &= \sum_{i=1}^{k} \sum_{j=k+1}^{n+1} D_{ij} \\ &= \sum_{i=1}^{k} \left[ \sum_{j=k+1}^{n} D_{ij} + D_{i,n+1} \right] \\ &= U_{k,n} + \sum_{i=1}^{k} D_{i,n+1}. \end{aligned}$$

Thus, each $U_{k,n+1}$ can be calculated from $U_{k,n}$ using just two additions—one to update the running total $\sum_{i=1}^{k} D_{i,n+1} = \sum_{i=1}^{k-1} D_{i,n+1} + D_{k,n+1}$ and one to add this running total to $U_{k,n}$, for a total of $2n+1$ operations.

Computing $T_{k,n}$ from $U_{k,n}$ requires an additional four operations, making a total of $O(6n)$ arithmetic operations and $n$ comparisons. Thus, somewhat surprisingly, the nonparametric formulation not only also scales linearly, but in fact turns out to be computationally faster than the parametric and to have the same storage requirements.

## Evaluation of Performance

The performance of the chart is measured in terms of its ARL following a step change in distribution. Apart from the in-control ARL, two factors influence the performance—the magnitude of the shift and the length of time for which the process runs in control before the step change. The relevance of this second factor may be unintuitive initially, but as the change-point formulation involves two-sample comparisons, its performance is affected by the size of the in-control data set as well as that of the run of out-of-control data.

In addition, it is of interest to know how the performance of the nonparametric chart compares with that of the parametric HQK chart in circumstances where both would be appropriate. A further performance comparison with the Zhou et al. (2009) method is also indicated to evaluate the impact of the smoothing they suggest for the $T_{k,n}$.

We explored the first two issues by simulating three settings: a step shift occurring at the 15th, 50th, and 500th observations, which means the true change point of $\tau = 14$, $\tau = 49$, or $\tau = 499$. For each combination of $\tau$ and $\delta$, we simulated 200,000 sequences. Testing started at the 15th process reading and used control limits giving an in-control ARL of 500 ($\alpha = 0.002$.) The necessary control limits were downloaded from the Web site www.stat.umn/edu/hawkins.

The process readings simulated were $N(0,1)$ up to time $\tau$ and $N(\delta,1)$ after time $\tau$, and so are amenable to both the parametric HQK and the nonparametric approach. Figure 1 presents the resulting ARLs for $\delta$ values in the range of $[0,3]$. The numeric values of the ARLs are given in Table 3; the values listed have a standard error of 0.2%.

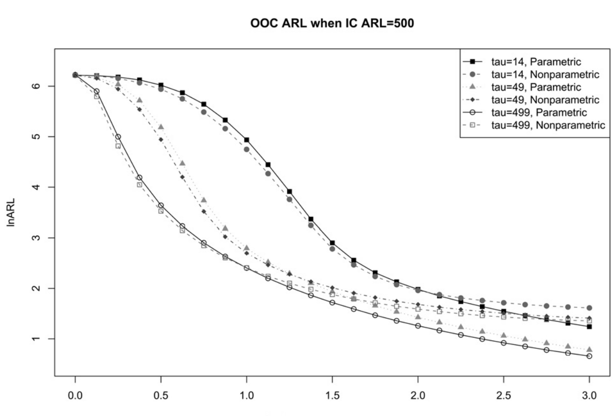


FIGURE 1. Performance of Parametric and Nonparametric Approaches with Different True Change Points.

TABLE 3. Comparing Parametric and Nonparametric

| $\tau$ | 14 | | 49 | | 499 | |
|---|---|---|---|---|---|---|
| $\delta$ | Par | Nonp | Par | Nonp | Par | Nonp |
| 0.125 | 497.09 | 493.52 | 480.66 | 471.60 | 366.10 | 328.55 |
| 0.250 | 483.66 | 471.51 | 417.86 | 381.19 | 148.38 | 123.52 |
| 0.375 | 456.91 | 431.80 | 303.75 | 254.45 | 66.09 | 57.30 |
| 0.500 | 412.01 | 380.67 | 178.50 | 140.06 | 38.09 | 33.93 |
| 0.625 | 354.48 | 314.36 | 86.92 | 66.71 | 25.30 | 23.20 |
| 0.750 | 282.55 | 241.91 | 41.92 | 33.72 | 18.26 | 17.22 |
| 0.875 | 206.27 | 174.13 | 24.05 | 20.53 | 13.91 | 13.55 |
| 1.000 | 139.35 | 115.43 | 16.34 | 14.84 | 11.07 | 11.11 |
| 1.125 | 85.24 | 71.46 | 12.39 | 11.75 | 9.02 | 9.38 |
| 1.250 | 50.09 | 42.89 | 9.90 | 9.78 | 7.55 | 8.19 |
| 1.375 | 28.98 | 25.61 | 8.20 | 8.44 | 6.44 | 7.25 |
| 1.500 | 18.20 | 16.17 | 6.93 | 7.48 | 5.58 | 6.54 |
| 1.625 | 12.91 | 11.79 | 5.99 | 6.73 | 4.91 | 5.99 |
| 1.750 | 10.07 | 9.37 | 5.25 | 6.16 | 4.34 | 5.55 |
| 1.875 | 8.41 | 7.96 | 4.65 | 5.73 | 3.89 | 5.18 |
| 2.000 | 7.24 | 7.10 | 4.16 | 5.38 | 3.52 | 4.90 |
| 2.125 | 6.36 | 6.53 | 3.76 | 5.09 | 3.21 | 4.67 |
| 2.250 | 5.69 | 6.11 | 3.42 | 4.86 | 2.94 | 4.47 |
| 2.375 | 5.15 | 5.81 | 3.13 | 4.67 | 2.71 | 4.31 |
| 2.500 | 4.69 | 5.56 | 2.89 | 4.51 | 2.51 | 4.19 |
| 2.625 | 4.31 | 5.36 | 2.68 | 4.37 | 2.34 | 4.08 |
| 2.750 | 3.98 | 5.22 | 2.49 | 4.26 | 2.18 | 4.00 |
| 2.875 | 3.71 | 5.11 | 2.33 | 4.17 | 2.05 | 3.93 |
| 3.000 | 3.46 | 5.02 | 2.18 | 4.10 | 1.93 | 3.88 |

Figure 1 shows some expected features. Having 49 rather than 14 initial in-control readings speeds the detection of a shift. The benefit is greatest when the shift is small. It is also interesting that having 499 initial in-control readings still gives a considerable improvement over having 49 initial readings when the shift is small.

A more interesting difference is between parametric and nonparametric. It is natural to expect that the nonparametric change-point model could beat the parametric for a nonnormal, especially skewed, underlying distribution. When the true distribution is normal, however, one would expect the parametric method to outperform the nonparametric across the board, though substantially so only for large shifts. Surprisingly, however, the nonparametric formulation offers you more than that. For $\tau = 14$, the nonparametric procedure has the faster reaction for all shifts up to 2.1 standard deviations, and only for larger shifts does the anticipated superiority of the parametric procedure surface.

For $\tau = 49$, the same behavior is seen, except that the crossover point between the two methodologies occurs at 1.5 standard deviations. And for $\tau = 499$, the two curves cross at 1 standard deviation. At all three, for small to moderate shifts, the nonparametric change-point model outperforms the parametric even when the underlying distribution is normal and the parametric model assumptions hold. For large shifts, intuition is right and the parametric approach is better.

Though initially surprising, the performance difference has an explanation. The parametric test is well known in the null case to have a strong tendency to split off just one or two observations from one end of the series. The nonparametric test is much less inclined to do so. This leads to the situation seen in Table 2 that, despite the identical correlation structure and largely matching distribution, the control limits of the nonparametric procedure are substantially smaller than those of the parametric. This lower threshold for a chart signal then allows the nonparametric procedure to react faster to moderate shifts. As usual, the parametric approach is better for large shifts.

The other relevant comparison is with the smoothed approach of Zhou et al (2009). To explore this, we simulated sequences with 150 in-control observations, followed by a shift. Both procedures used 50 initial warm-up observations prior to the start of testing and an in-control ARL of 500. The resulting out-of-control ARLs are shown in Table 4, the run lengths having standard errors of 1%.

It appears that the additional complication of the smoothing adds a modest benefit for small shifts, as the Zhou et al. (2009) procedure has a somewhat faster response for small shifts (14% shorter ARL for a shift of 0.375 $\sigma$). However, it leads to slower response for shifts above $0.75\sigma$, where the run length is up to 25% longer.

## Example

Aluminum smelters monitor the feed material to respond to changes in its composition. One of the nuisance constituents is silica. Table 5 lists a data set (kindly provided by Len Homer) on the silica concentrations of the feed to a smelter. The sequence is also plotted in Figure 2. A glance at the figure is enough to show that the numbers do not follow a normal distribution; there is a clustering of values close to the axis, with isolated values far above. The graph also

TABLE 4. Comparing Smoothed and Unsmoothed

| $\delta$ | Zhou | Current |
|---|---|---|
| 0.125 | 394.29 | 419.55 |
| 0.250 | 197.91 | 225.69 |
| 0.375 | 79.65 | 92.72 |
| 0.500 | 39.77 | 43.85 |
| 0.625 | 25.59 | 26.68 |
| 0.750 | 18.80 | 18.95 |
| 0.875 | 14.78 | 14.53 |
| 1.000 | 12.21 | 11.66 |
| 1.125 | 10.45 | 9.78 |
| 1.250 | 9.17 | 8.37 |
| 1.375 | 8.22 | 7.41 |
| 1.500 | 7.51 | 6.70 |
| 1.625 | 6.93 | 6.10 |
| 1.750 | 6.52 | 5.60 |
| 1.875 | 6.15 | 5.25 |
| 2.000 | 5.86 | 4.95 |
| 2.125 | 5.62 | 4.70 |
| 2.250 | 5.43 | 4.50 |
| 2.375 | 5.27 | 4.33 |
| 2.500 | 5.14 | 4.20 |
| 2.625 | 5.04 | 4.10 |
| 2.750 | 4.96 | 4.00 |
| 2.875 | 4.89 | 3.92 |
| 3.000 | 4.82 | 3.87 |

suggests that the values at the end are higher than those at the beginning, but it is less clear when the shift might have occurred or how soon thereafter it could have been claimed confidently.

One possible approach is to try finding a transformation to normality, following which the parametric HQK procedure could be applied. This is complicated by the fact that the sequence is relatively short and compounded by the fact that if there is a change point, then the data set is a mixture rather than a single distribution. Nevertheless, a log transformation seems reasonable and is supported by the general observation that low-concentration analytes tend to follow log-normal distributions. We can then apply the parametric HQK procedure to the log-transformed values.

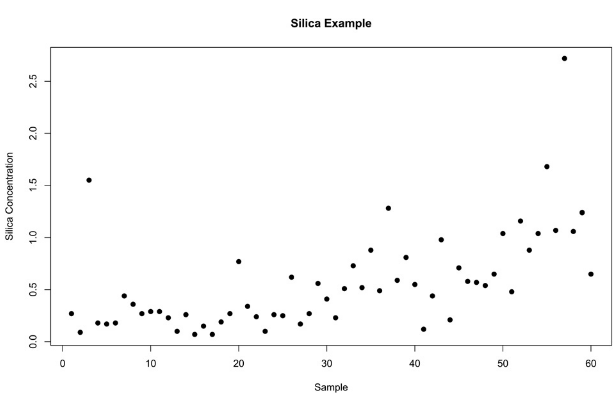


FIGURE 2. Time-Ordered Plot of SiO2 Data.

TABLE 5. Silica Concentration in Smelter

| $n$ | SiO2 | $n$ | SiO2 | $n$ | SiO2 | $n$ | SiO2 |
|---|---|---|---|---|---|---|---|
| 1 | 0.27 | 16 | 0.15 | 31 | 0.23 | 46 | 0.58 |
| 2 | 0.09 | 17 | 0.07 | 32 | 0.51 | 47 | 0.57 |
| 3 | 1.55 | 18 | 0.19 | 33 | 0.73 | 48 | 0.54 |
| 4 | 0.18 | 19 | 0.27 | 34 | 0.52 | 49 | 0.65 |
| 5 | 0.17 | 20 | 0.77 | 35 | 0.88 | 50 | 1.04 |
| 6 | 0.18 | 21 | 0.34 | 36 | 0.49 | 51 | 0.48 |
| 7 | 0.44 | 22 | 0.24 | 37 | 1.28 | 52 | 1.16 |
| 8 | 0.36 | 23 | 0.10 | 38 | 0.59 | 53 | 0.88 |
| 9 | 0.27 | 24 | 0.26 | 39 | 0.81 | 54 | 1.04 |
| 10 | 0.29 | 25 | 0.25 | 40 | 0.55 | 55 | 1.68 |
| 11 | 0.29 | 26 | 0.62 | 41 | 0.12 | 56 | 1.07 |
| 12 | 0.23 | 27 | 0.17 | 42 | 0.44 | 57 | 2.72 |
| 13 | 0.10 | 28 | 0.27 | 43 | 0.98 | 58 | 1.06 |
| 14 | 0.26 | 29 | 0.56 | 44 | 0.21 | 59 | 1.24 |
| 15 | 0.07 | 30 | 0.41 | 45 | 0.71 | 60 | 0.65 |

The nonparametric change-point model can be applied immediately to the data on the natural scale.

The results are shown in Figure 3 and Figure 4, respectively. In each figure, the left panel shows the change-point statistic along with the control limit for an in-control ARL of 500. The right panel shows the maximizing $k$—the estimate of the last in-control reading.

Both nonparametric and parametric control charts conclude that there is a location shift midway through the series. The nonparametric chart exceeds its control limit at $n = 37$ and remains above the control limit to the end of the data. The parametric chart signals a bit later, at $n = 39$, and then dips back below the control limit before re-emerging at process reading $n = 44$.

As for the estimate of the change point, both methods vacillate initially between $\hat{\tau} = 31$ and $\hat{\tau} = 28$ before settling on the former.

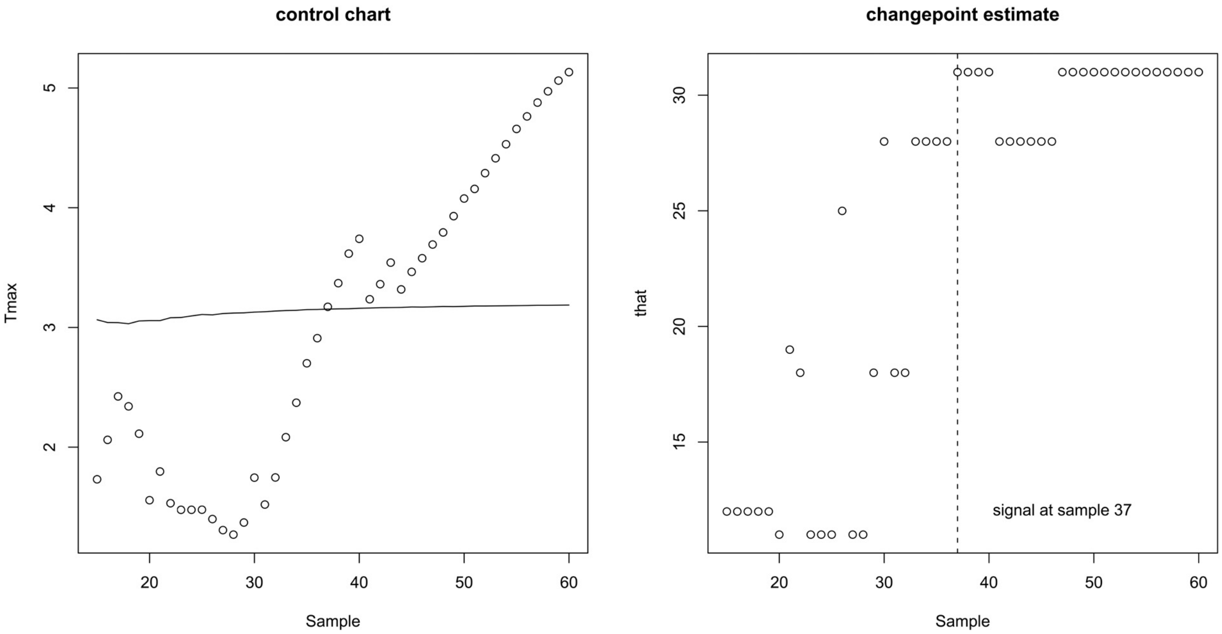


FIGURE 3. Nonparametric Change-Point Model on Original SiO2 Data: $T_{\max,n}$ and $h_{n,0.002}$.

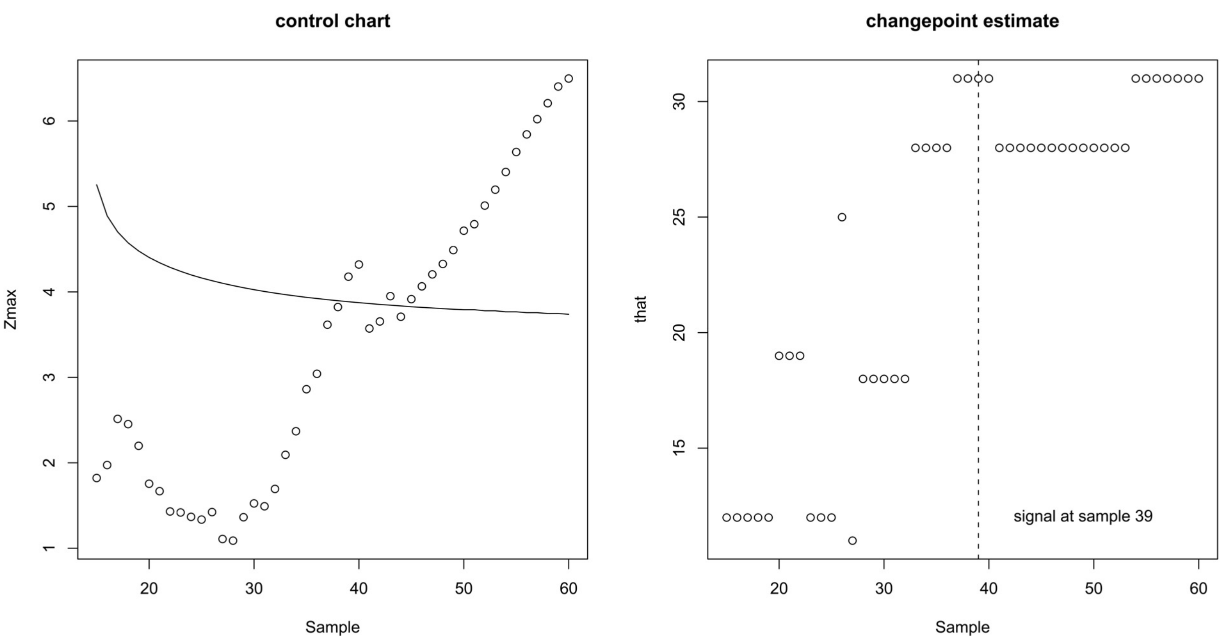


FIGURE 4. Parametric Change-Point Model on Logged SiO2 Data: $Z_{\max,n}$ and $h_{n,0.002}$.

In this example then, the nonparametric change-point control chart provides results that are compatible with those of the parametric method but is a little faster. These apparent advantages come on top of the fact that there is no need to struggle with searching for the appropriate transformation and worrying about the distribution issue at all.

## Conclusion

The introduction of the change-point model into Phase II statistical process control has brought a new perspective to the problem of unknown parameters, and our nonparametric change-point model takes this a step further. In statistical theory, a simple two-sample Mann–Whitney test has a relative efficiency as high as 96% when the distribution is normal, while it is a great improvement when the true distribution is not. Our results show something even better—that, for moderate shifts, there appears to be a gain rather than a loss of performance in using a nonparametric rather than a parametric approach. Moreover, it requires minimal assumptions and knowledge of baseline data to detect a change in an ongoing process, and can give the greatest freedom from worry about the underlying distribution. This method extends the applicability of previous proposals for Phase II change-point models.

## Acknowledgments

The authors are grateful to the editor and referees for a number of suggestions for improvement.

## References

Amin, R. W.; Reynolds, M. R.; and Bakir, S. T. (1995). "Nonparametric Quality Control Charts Based on the Sign Statistic". *Communications in Statistics—Theory and Methods* 24, pp. 1597–1623.

Amin, R. and Search, A. J. (1991). "A Nonparametric Exponentially Weighted Moving Average Control Scheme". *Communications in Statistics—Simulation and Computation* 20, pp. 1049–1072.

Bakir, S. T. and Reynolds, M. R. (1979). "A Nonparametric Procedure for Process Control Based on Within-Group Ranking". *Technometrics* 21, pp. 175–183.

Chakraborti, S.; Van Der Laan, P.; and Bakir, S. T. (2001). "Nonparametric Control Charts: An Overview and Some Results". *Journal of Quality Technology* 33(3), pp. 304–315.

Conover, W. J. (1999). *Practical Nonparametric Statistics*. New York, NY: Wiley.

Csörgő, M. and Horváth, L. (1997). *Limited Theorems in Change-Point Analysis*. Chichester: Wiley.

Deng, Q. (2009). "A Nonparametric Change-Point Model for Phase II Analysis". Ph.D. thesis, School of Statistics, University of Minnesota.

Janacek, G. J. and Meikle, S. E. (1997). "Control Charts Based on Medians". *The Statistician* 46, pp. 19–31.

Jensen, W. A.; Jones-Farmer, L. A.; Champ, C. W.; and Woodall, W. H. (2006). "Effects of Parameter Estimation on Control Chart Properties: A Literature Review". *Journal of Quality Technology* 38, pp. 349–364.

Jones, L. A.; Champ, C. W.; and Rigdon, S. E. (2001). "The Performance of Exponentially Weighted Moving Average Charts with Estimated Parameters". *Technometrics* 43, pp. 156–167.

Jones, L. A.; Champ, C. W.; and Rigdon, S. E. (2004). "The Run Length Distribution of the CUSUM with Estimated Parameters". *Journal of Quality Technology* 36, pp. 95-108.

Hackl, P. and Ledolter, J. (1991). "A Control Chart Based on Ranks". *Journal of Quality Technology* 23, pp. 117–124.

Hawkins, D. M. (1977). "Testing a Sequence of Observations for a Shift in Location". *Journal of the American Statistical Association* 72(357), pp. 180–186.

Hawkins, D. M.; Qiu, P.; and Kang, C. W. (2002). "The Changepoint Model for Statistical Process Control". *Journal of Quality Technology*, 35, 355-365.

Hinkley, D. V. (1970). "Inference about the Change-Point in a Sequence of Random Variables". *Biometrika* 57(1), pp. 1–17.

Irvine, J. M. (1982). "Changes in Regime in Regression Models". Ph.D. dissertation, Yale University.

Jones-Farmer, L. A.; Jordan, V.; and Champ, C. W. (2009). "Distribution-Free Phase I Control Charts for Subgroup Location". *Journal of Quality Technology* 41, pp. 304–316.

Lai, T. L. (2001). "Sequential Analysis: Some Classical Problems and New Challenges". *Statistica Sinica*, 11, pp. 303–408.

McDonald, D. (1990). "A CUSUM Procedure Based on Sequential Ranks". *Naval Research Logistics* 37, pp. 627–646.

Pettitt, A. N. (1979). "A Non-Parametric Approach to the Change-Point Problem. *Applied Statistics* 28(2), pp. 126–135.

Schechtman, E. and Wolfe, D. A. (1981). "Distribution-Free Tests for the Changepoint Problem". Technical Report, The Ohio State University.

Schechtman, E. (1982). "A Nonparametric Test for Detecting Changes in Location". *Communications in Statistics—Theory and Methods* 11(13), pp. 1475–1482.

Schechtman, E. and Wolfe, D. A. (1984). "Nonparametric Statistical Procedures for the Changepoint Problem". *Journal of Statistical Planning and Inference* 9, pp. 389–396.

Siegmund, D. and Venkatraman, E. S. (1995). "Using the Generalized Likelihood Ratio Statistic for Sequential Detection of a Change-Point". *The Annals of Statistics* 23(1), pp. 255–271.

Willemain, T. R. and Runger, G. C. (1996). "Designing Control Charts Using an Empirical Reference Distribution". *Journal of Quality Technology* 28, pp. 31–38.

Woodall, W. H. (2000). "Controversies and Contradictions in Statistical Process Control (with Discussion)". *Journal of Quality Technology* 32, pp. 341–378.

Woodall, W. H. and Montgomery, D. C. (1999). "Research Issues and Ideas in Statistical Process Control". *Journal of Quality Technology* 31, pp. 376–386.

Zou, C. and Tsung, F. (2010). "Likelihood Ratio Based Distribution-Free EWMA Control Charts". *Journal of Quality Technology* 42, pp. 174–196.

Zhou, C.; Zou, C.; Zhang, Y.; and Wang, Z. (2009). "Nonparametric Control Chart Based on Changepoint Model". *Statistical Papers* 50, pp. 13–28.