Research Article

Quality and Reliability Engineering International

(wileyonlinelibrary.com) DOI: 10.1002/qre.1835

Published online 7 July 2015 in Wiley Online Library

A comparative study of memory-type control charts under normal and contaminated normal environments

Hafiz Zafar Nazir,[a*†] Nasir Abbas,[b] Muhammad Riaz[b] and Ronald J.M.M. Does[c]

Cumulative sum (CUSUM) and exponentially weighted moving average (EWMA) control charts are commonly used to detect small changes in the parameters of production processes. Recently, a new control structure was introduced, named as mixed EWMA–CUSUM control chart, which combined both charts. The current study provides a detailed comparison of these three types of control charts when the process parameters are unknown under normal and contaminated normal environments. Performance measures average run length and different percentiles of run length distribution are used for comparison purposes. We investigate six different location estimators with the structures of the three memory charts and study their robustness properties. Copyright © 2015 John Wiley & Sons, Ltd.

Keywords: memory type charts; average run length; percentiles; robust estimators; contaminated environment; diffuse symmetric disturbances; efficiency

1. Introduction

Control charts may be classified in memory-less (Shewhart-type) and memory control charts. Shewhart-type charts use current available information and are less sensitive to detect small and moderate changes in the process parameters, but are most efficient at detecting large shifts. An approach to deal with the detection of small shifts is to use memory control charts, such as the cumulative sum (CUSUM) control chart proposed by Page[1] and the exponentially weighted moving average (EWMA) control chart proposed by Roberts.[2] These charts are designed such that they use the past information along with the current information, which makes them very sensitive to shifts of small and moderate magnitudes in the process parameters. A number of modifications of the CUSUM and EWMA charts have been developed to further enhance the performance of these charts. Some of these enhancements may be seen in Lucas,[3] Lucas and Saccucci,[4] Steiner,[5] Capizzi and Masarotto,[6] Zhao et al.,[7] Colosimo et al.,[8] Castagliola et al.,[9] Machado and Costa,[10] Riaz et al.,[11] Abbas et al.[12] and the references therein. Following these authors, Abbas et al.[13] proposed a mixed EWMA–CUSUM control chart and concluded that mixing the two charts makes the proposed scheme even more sensitive to small shifts in the process mean as compared to the other schemes designed for similar purposes.

In practice, process parameters are unknown, and they need to be estimated from samples, which are assumed to be in state of statistical control. Woodall and Montgomery[14] name this stage as Phase I. The resultant estimates from Phase I establish the control limits that are used to monitor the process parameter of interest in the next stage: Phase II.

Jensen et al.[15] studied estimation effects on control chart properties in Phase I and their Phase II impact. Schoonhoven et al.[16] used different robust estimators for the location control chart in a Shewhart set up by considering limited information in Phase I and studied the performance of these estimators in Phase II. Recently, Nazir et al.[17] proposed the use of some robust location estimators with the control structures of the CUSUM chart, in order to increase the robustness of the CUSUM chart against contamination or non-normality. However, they only considered the situation when a large number of samples are available in Phase I, and they did not take into account the estimation effects of parameters.

The concern of this study is to assess the estimation effects of process parameters in Phase I and to check the impact and influence of robust estimators of the location parameter on the Phase II performance with the design structures of CUSUM, EWMA and mixed EWMA–CUSUM control charts under different environments. Generally, the performance and efficiency of control charts are assessed by the determinant, called average run length (*ARL*). The *ARL* is the mean of a random variable called

[a]*Department of Statistics, University of Sargodha, Sargodha, Pakistan*
[b]*Department of Mathematics and Statistics, King Fahad University of Petroleum and Minerals, Dhahran, Saudi Arabia*
[c]*Department of Operations Management IBIS UvA, University of Amsterdam, Plantage Muidergracht 12, 1018 TV, Amsterdam, The Netherlands*
**Correspondence to: Hafiz Zafar Nazir, Department of Statistics, University of Sargodha, Pakistan.*
[†]*E-mail: hafizzafarnazir@yahoo.com*

run length (RL), where RL is the number of samples required before an alarm signal occurs. The in-control *ARL* is likely to be high, but can be fixed to a specific number for a given false alarm rate and is denoted by ARL_0. The out-of-control *ARL* is expected to be as small as possible and is denoted by ARL_1. Besides using the *ARLs* as efficiency indicators, we will also take into account the standard deviation of the run length (SDRL) and different percentiles of the *RL* distribution in order to have an even better comparison.

In the next section, we give the details regarding the control structure of three memory-type control charts and the Phase I and Phase II estimators. Design structures of these charts are provided in section . In section we provide a comprehensive comparison of the three memory-type control charts (based on six different location estimates) in terms of *ARL*s and percentiles of *RL* distribution to study their robustness. Finally, the article is concluded in section 5.

2. CUSUM, EWMA and mixed EWMA–CUSUM charts

Shewhart-type control charts are less efficient to detect small and moderate shifts in the process parameter(s). For that reason, some memory-type control charts are proposed. The most important ones include the CUSUM, EWMA and mixed EWMA–CUSUM charts, and the current section contains the details about these three structures.

2.1. Cumulative sum charts

Page[1] presented the idea of accumulating the positive and negative deviations from the process location in two different statistics C_i^+ and C_i^-, respectively. These two statistics are defined as:

$$C_i^+ = \max\left[0, \left(\hat{\theta}_i - \theta_0\right) - K_{\hat{\theta}} + C_{i-1}^+\right] \ , \ C_i^- = \max\left[0, -\left(\hat{\theta}_i - \theta_0\right) - K_{\hat{\theta}} + C_{i-1}^-\right] \tag{1}$$

where i is the sample number and $\hat{\theta}$ is the location estimator used to monitor the process location parameter. The initial values for both of the statistics given in (1) are usually taken equal to the target value θ_0, i.e. $C_0^+ = C_0^- = \theta_0$. The statistics C_i^+ and C_i^- are plotted against the control limit $H_{\hat{\theta}}$, and an out-of-control signal is generated if either one of these statistics crosses the control limit. The standardized versions of the chart parameters ($K_{\hat{\theta}}$ and $H_{\hat{\theta}}$) are given as:

$$K_{\hat{\theta}} = k \times \sigma_{\hat{\theta}} \ , \ H_{\hat{\theta}} = h \times \sigma_{\hat{\theta}}. \tag{2}$$

Here *k* and *h* are the constants which are chosen to satisfy a pre-specified ARL_0.

2.2. Exponentially weighted moving average charts

Roberts[2] proposed a control charting scheme, in which the plotting statistic is split into two components (i.e. present information and past information), and named it as the exponentially weighted moving average (EWMA) chart. The weights are assigned to the observations such that these weights decrease exponentially for the mor e dated observations. The control structure of the EWMA chart, consisting of a plotting statistic and the control limits, is given as:

$$Z_i = \lambda\hat{\theta}_i + (1-\lambda)Z_{i-1} \tag{3}$$

$$\left.LCL_i = \theta_0 - L_{\hat{\theta}}\sqrt{\text{Var}\left(\hat{\theta}\right) \times \frac{\lambda}{2-\lambda}\left(1-(1-\lambda)^{2i}\right)} CL = \theta_0 UCL_i = \theta_0 + L_{\hat{\theta}}\sqrt{\text{Var}\left(\hat{\theta}\right) \times \frac{\lambda}{2-\lambda}\left(1-(1-\lambda)^{2i}\right)}\right\} \tag{4}$$

where $\lambda \in (0, 1]$ is the smoothing parameter of the chart. The initial value for the above plotting statistic in (3) is usually taken equal to the target value, i.e. $Z_0 = \theta_0$. $L_{\hat{\theta}}$ is the control limit coefficient and can be chosen to satisfy the pre-specified ARL_0. Note that for $\lambda = 1$, we obtain the Shewhart control chart, and hence the Shewhart control chart is a special case of the EWMA control chart.

2.3. Mixed EWMA–CUSUM charts

Abbas et al.[13] proposed a mixture of the CUSUM and EWMA charts and named it as the mixed EWMA–CUSUM chart. The two plotting statistics (M_i^+ and M_i^-) for this chart are given as:

$$M_i^+ = \max\left[0, (Z_i - \theta_0) - A_{\hat{\theta},i} + M_{i-1}^+\right] \ , \ M_i^- = \max\left[0, -(Z_i - \theta_0) - A_{\hat{\theta},i} + M_{i-1}^-\right] \tag{5}$$

where Z_i is defined as in (3). The statistics (given in (5)) are plotted against the control limit $B_{\hat{\theta},i}$ and an out-of-control signal is detected if either one of these statistics crosses the control limit. The standardized versions $A_{\hat{\theta},i}$ and $B_{\hat{\theta},i}$ are given as:

$$A_{\hat{\theta},i} = a \times \sqrt{\text{Var}\left(\hat{\theta}\right) \times \frac{\lambda}{2-\lambda}\left(1-(1-\lambda)^{2i}\right)} \ , \ B_{\hat{\theta},i} = b \times \sqrt{\text{Var}\left(\hat{\theta}\right) \times \frac{\lambda}{2-\lambda}\left(1-(1-\lambda)^{2i}\right)} \tag{6}$$

where a and *b* are constants like *k* and *h* in (2), respectively.

Previous equations (1)–(6) include θ_0 and $\text{Var}\left(\hat{\theta}\right)$. Consider that X_{ij}, $i=1,2,\ldots,n$ and $j=1,2,\ldots,m$ denote the Phase I data, when the process is in an in-control state and let Y_{ij}, $i=1,2,\ldots,n$ and $j=1,2,\ldots,$ denote the Phase II data.

We assume that the X_{ij} are normally distributed with mean θ_0 and variance σ^2, i.e. $N(\theta_0,\sigma^2)$. The unknown location parameter θ_0 is estimated from the mean of the sample means, i.e.

$$\hat{\theta}_0=\overline{\overline{X}}=\frac{1}{m}\sum_{j=1}^{m}\overline{X}_j=\frac{1}{m}\sum_{j=1}^{m}\left(\frac{1}{n}\sum_{i=1}^{n}X_{ij}\right)$$

and the unknown dispersion parameter σ is based on the pooled sample standard deviation,

$$S_p=\left(\frac{1}{m}\sum_{j=1}^{m}S_j^2\right)^{1/2}$$

where S_j^2 *is the jth* sample variance defined by

$$S_j^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(X_{ij}-\overline{X}_j\right)^2.$$

An unbiased estimator of σ is given by $\hat{\sigma}=S_p/c_4(m(n-1)+1)$, where $c_4(q)$ is defined by

$$c_4(q)=\left(\frac{2}{q-1}\right)^{1/2}\frac{\Gamma\left(\frac{q}{2}\right)}{\Gamma\left(\frac{q-1}{2}\right)}.$$

Note that $\text{Var}\left(\hat{\theta}\right)$ is a function of dispersion parameter σ which is unknown and has to be estimated. The estimates $\hat{\theta}_0=\overline{\overline{X}}$ and $\hat{\sigma}=S_p/c_4(m(n-1)+1)$ will be used in (1)–(6), respectively, instead of θ_0 and σ for the construction of the control limits in Phase I.

In Phase II we assume that Y_{ij} are independent and equally distributed as the X_{ij}, with the only difference that the location parameter may be shifted.

The sample mean $\overline{Y}=\frac{1}{n}\sum_{i=1}^{n}Y_i$ is one of the estimators of the population location that can replace $\hat{\theta}$ in (1)–(6) in Phase II. However, there are many other estimators that can also be used instead of $\hat{\theta}$ with the above mentioned three memory charts structures. The first estimator we consider out of those is the sample median. The sample median is defined as the middle order statistic $\tilde{Y}=Y_{\left(\frac{n+1}{2}\right)}$ for odd sample sizes and the average of the two middle order statistics $\tilde{Y}=\frac{1}{2}\left(Y_{\left(\frac{n}{2}\right)}+Y_{\left(\frac{n+2}{2}\right)}\right)$ in case of even sample sizes. The sample median is a robust estimator, because it is least affected by outliers (cf. Dixon and Massey[18]). The next estimator is the sample midrange and is defined as $MR=\frac{Y_{(1)}+Y_{(n)}}{2}$, where $Y_{(1)}$ and $Y_{(n)}$ are the lowest and highest order statistics in a random sample of size n. It is highly sensitive to outliers as its design structure is based on only extreme values of data (cf. Ferrell[19] for more details). We also include the estimator based on the median of the pairwise Walsh averages, which is defined as: $HL=\text{median}((Y_j+Y_k)/2, 1\le j\le k\le n)$. The main advantage of the HL estimator is that it is robust against outliers in a sample. For more properties of HL see Hettmansperger and McKean.[20] The estimator HL is also known as the Hodges–Lehmann estimator. The next estimator included in this study is the trimean of a sample, which is the weighted average of the sample median and two quartiles and is defined as: $TM=\frac{Q_1+2Q_2+Q_3}{4}$, where Q_p $(p=1,2,3)$ denote one of the three quartiles in a sample. For detailed properties of trimean (TM) see Wang et al..[21] The last estimator used in this study is sample trimmed mean and is defined as $T_RM=\frac{1}{n-2T}\sum_{i=T+1}^{n-T}Y_{(i)}$, where $2T$ is the number of trimmed values and $Y_{(i)}$ is the i^{th} order statistic in a sample of size n. We take $T=1$ for $n=5$ and $T=2$ for $n=10$, respectively.

Under normality, the means and (asymptotic) variances (cf. Song et al.,[22] Caperaa and Rivest,[23] Khattree and Rao[24] and Wang et al.[21]) of these estimators are given in Table I.

3. Design and derivation of phase II control limits of the charts

The design of the Phase II control charts involves a derivation of different factors: the CUSUM structure requires values of k and h (cf. (2)), the EWMA scheme needs λ and L (cf. (4)), and the mixed EWMA–CUSUM demands values of a and b (cf. (6)) for the construction of the control limits of these charts. Along with these factors, the Phase I process location parameter θ and dispersion parameter σ also have to be estimated.

We derive these factors in such a way that we obtain the intended value of $ARL_0=370$. The Phase I estimators are $\hat{\theta}_0=\overline{\overline{X}}$ and $\hat{\sigma}=S_p/c_4(m(n-1)+1)$. We employ different estimators for the Phase II plotting statistic(s) and adopt the following settings, $k=0.5$, $\lambda=0.13$ and $a=0.5$ as optimal constants to detect a shift size of $\delta=1$ (i.e. a shift of 1 σ), for respectively, CUSUM, EWMA and mixed EWMA–CUSUM charts, taking inspiration from Lucas,[3] Crowder,[25] and Abbas et al..[13] We simulate the factors h for the CUSUM, L for the EWMA and b for the mixed EWMA–CUSUM control chart by considering $m=50$ subgroups of sizes $n=5$ and $n=10$ from an uncontaminated normal environment with desired $ARL_0=370$. Values of these factors are given in Table II.

Table I. Expected values and the asymptotic variances of the estimators under in-control situation

Estimator	Expected value of the estimator	(Asymptotic) variance of the estimator
Mean	θ_0	$\frac{\sigma^2}{n}$
Median	θ_0	$\frac{\pi\sigma^2}{2n}$
Midrange	θ_0	$\frac{\pi^2\sigma^2}{24\ \ln(n)}$
Hodges–Lehmann (HL)	θ_0	$\frac{\pi\sigma^2}{3n}$
Trimean	θ_0	$\frac{\pi\sigma^2}{2.61n}$
Trimmed mean	θ_0	$\frac{n(\sigma^2-2.9565)}{(n-2)^2}$ for $n=5$ $\frac{n(\sigma^2-5.9209)}{(n-4)^2}$ for $n=10$

Table II. Factors of CUSUM, EWMA and mixed EWMA–CUSUM charts under uncontaminated normal environment with $\boldsymbol{m}=50$ at $\boldsymbol{ARL}_0=370$

		Phase II estimators					
n	Chart	Mean	Median	Midrange	HL	Trimean	Trimmed
5	CUSUM	$h=5.000$	$h=4.532$	$h=5.058$	$h=5.110$	$h=4.480$	$h=4.520$
	EWMA	$L=2.895$	$L=2.740$	$L=2.912$	$L=2.926$	$L=2.790$	$L=2.760$
	Mixed	$b=36.30$	$b=31.76$	$b=35.90$	$b=36.82$	$b=32.20$	$b=33.98$
10	CUSUM	$h=5.078$	$h=4.440$	$h=5.168$	$h=5.116$	$h=4.680$	$h=4.827$
	EWMA	$L=2.916$	$L=2.714$	$L=2.935$	$L=2.929$	$L=2.790$	$L=2.847$
	Mixed	$b=36.70$	$b=31.08$	$b=35.26$	$b=36.78$	$b=33.50$	$b=35.14$

4. Performance evaluation of memory control charts

This section gives the details regarding the performance evaluation of the three memory control charts for normal and contaminated normal environments.

The performance of the design structures is measured in terms of different characteristics of run length distribution e.g. its *ARL* s, SDRL and its various percentile points. As a baseline we use the most conventional measure in-control *ARL*, i.e. $ARL_0=370$, under normality for the performance evaluation and comparisons. The evaluation of *ARL* values is done using Monte Carlo simulations.

Now for the Phase II analysis, the estimated Phase I process location parameter θ and dispersion parameter σ are used for constructing the control limits of all the charts (where the control limits of the CUSUM are given in (2), the control limits of the EWMA are given in (4) and the control limits of the mixed EWMA–CUSUM control chart are given in (6)). Then, by applying the out-of-control condition (i.e. when an out-of-control signal occurs), we have noted the sample number where the plotting statistic crosses the control limits. This noted number is called run length which is replicated 10^5 times to get the run length distribution. The mean of that distribution, when the location has not been changed, is known as ARL_0. After that, we introduce different amounts of shifts δ (i.e. the location parameter is shifted from the target value θ_0 to $\theta_0+\delta\sigma$ such that when $\delta=0$, the location parameter θ of the process is in control; otherwise the location parameter has changed and needs to be detected) in the process while keeping the control limits the same as we have used for the in-control case. It results in an evaluation of the ARL_1 performance of all the charts. Hence the performance measure ARL_1 is used for measuring the efficiency of charts and for robustness comparison, ARL_0 is used.

4.1. Normal and contaminated normal environments

The description of the environments, for which the performance of the CUSUM, EWMA and mixed EWMA–CUSUM control charts is evaluated, is given as follows:

4.1.1. Normal environment Here, we provide the performance of the memory charts under a perfectly normal environment with mean θ and variance σ^2, i.e. $N(\theta,\sigma^2)$. Without loss of generality, we use $\theta=0$ and $\sigma=1$ throughout this article.

4.1.2. Diffuse symmetric variance contaminated normal environment Here, $(1-\alpha)100\%$ observations in a sample come from the standard normal distribution, i.e. $N(0,1)$, and $(\alpha)100\%$ observations of that sample are from $N(0,4)$, i.e. a normal distribution with inflated variance.

4.1.3. Localized variance contaminated normal environment Here, a sample of size n with probability $(1-\alpha)100\%$ come from the standard normal distribution, i.e. $N(0,1)$, and otherwise a sample with probability $(\alpha)100\%$ is from $N(0,4)$, i.e. a normal distribution with inflated variance.

4.1.4. Diffuse asymmetric variance contaminated normal environment Here, asymmetric variance disturbances are introduced in the process, i.e. $(1-\alpha)100\%$ observations in a sample come from the standard normal distribution, i.e. $N(0,1)$, and $(\alpha)100\%$ observations are from $N(0,1)$ having a multiple of a $\chi^2_{(1)}$ variable added to it with multiplier equal to 4.

Note that these environments are commonly used in these robust studies (cf. Schoonhoven et al.[16] and Nazir et al.[17]). We take in the comparisons $\alpha = 0.05$.

Table III. ***ARL*** values of CUSUM, EWMA and mixed EWMA–CUSUM charts with $\boldsymbol{m} = 50$ at $\boldsymbol{ARL_0} = 370$ under uncontaminated normal environment

			ARL(δ)								
n	Chart	Estimator	0	0.25	0.5	0.75	1	1.5	2	3	5
5	CUSUM	Mean	370.960	39.833	9.073	5.043	3.551	2.322	1.861	1.124	1.000
		Median	370.186	56.332	12.110	6.264	4.264	2.682	2.055	1.374	1.000
		Midrange	369.489	50.029	11.147	5.973	4.118	2.633	2.039	1.358	1.000
		HL	371.060	42.524	9.614	5.314	3.714	2.413	1.920	1.188	1.000
		Trimean	371.919	43.767	9.487	5.125	3.571	2.317	1.841	1.116	1.000
		Trimmed	369.742	42.234	9.212	5.042	3.508	2.287	1.783	1.136	1.000
	EWMA	Mean	371.264	34.015	7.516	3.691	2.353	1.362	1.061	1.000	1.000
		Median	369.967	46.430	10.147	4.894	3.058	1.685	1.206	1.003	1.000
		Midrange	369.669	42.766	9.386	4.561	2.858	1.595	1.157	1.002	1.000
		HL	370.258	36.090	7.956	3.928	2.475	1.420	1.084	1.000	1.000
		Trimean	370.297	39.073	8.247	4.027	2.532	1.442	1.089	1.000	1.000
		Trimmed	369.247	35.086	7.694	3.828	2.433	1.412	1.087	1.000	1.000
	Mixed	Mean	369.937	37.119	17.361	12.508	10.101	7.589	6.219	4.773	3.179
	EWMA–	Median	369.702	43.617	19.256	13.650	10.950	8.169	6.675	5.057	3.630
	CUSUM	Midrange	369.226	42.226	19.098	13.625	10.960	8.195	6.710	5.084	3.692
		HL	369.097	38.443	17.893	12.861	10.386	7.788	6.381	4.889	3.342
		Trimean	370.070	37.442	17.238	12.363	9.978	7.468	6.121	4.668	3.103
		Trimmed	370.697	37.969	17.469	12.524	10.106	7.571	6.206	4.707	3.287
10	CUSUM	Mean	369.951	17.700	5.521	3.357	2.477	1.802	1.232	1.000	1.000
		Median	370.690	24.379	6.683	3.886	2.802	1.944	1.457	1.003	1.000
		Midrange	370.496	31.178	8.398	4.782	3.400	2.252	1.825	1.082	1.000
		HL	369.035	18.640	5.741	3.471	2.552	1.845	1.293	1.000	1.000
		Trimean	370.352	19.469	5.766	3.458	2.532	1.814	1.264	1.000	1.000
		Trimmed	370.627	18.988	5.735	3.450	2.537	1.801	1.283	1.001	1.000
	EWMA	Mean	371.352	15.161	4.110	2.170	1.471	1.035	1.000	1.000	1.000
		Median	370.566	20.177	5.304	2.745	1.794	1.132	1.007	1.000	1.000
		Midrange	370.449	26.687	6.751	3.412	2.191	1.291	1.041	1.000	1.000
		HL	369.956	15.943	4.316	2.270	1.524	1.046	1.001	1.000	1.000
		Trimean	368.507	16.440	4.421	2.313	1.554	1.056	1.001	1.000	1.000
		Trimmed	371.715	16.275	4.399	2.303	1.550	1.059	1.001	1.000	1.000
	Mixed	Mean	369.621	24.261	13.175	9.752	7.958	6.026	4.978	3.926	2.946
	EWMA–	Median	370.223	26.953	14.083	10.327	8.394	6.315	5.167	3.990	2.994
	CUSUM	Midrange	370.631	31.757	16.121	11.733	9.507	7.151	5.881	4.418	3.009
		HL	371.022	24.904	13.434	9.937	8.105	6.130	5.039	3.964	2.979
		Trimean	368.809	24.802	13.275	9.793	7.977	6.033	4.978	3.918	2.939
		Trimmed	370.225	25.018	13.421	9.898	8.070	6.097	5.010	3.874	2.864

Table IV. ***ARL*** values of CUSUM, EWMA and mixed EWMA–CUSUM charts with $\boldsymbol{m}=50$ at $\boldsymbol{ARL_0}=370$ under diffuse symmetric variance contaminated normal environment

			ARL(δ)								
n	Chart	Estimator	0	0.25	0.5	0.75	1	1.5	2	3	5
5	CUSUM	Mean	205.586	34.104	9.385	5.261	3.714	2.430	1.925	1.216	1.005
		Median	280.197	51.814	12.359	6.435	4.385	2.754	2.102	1.416	1.002
		Midrange	123.295	35.917	11.411	6.352	4.422	2.832	2.189	1.482	1.093
		HL	250.772	41.438	10.176	5.581	3.896	2.509	1.974	1.256	1.007
		Trimean	241.763	42.477	10.084	5.427	3.763	2.427	1.896	1.194	1.002
		Trimmed	258.276	40.950	9.651	5.241	3.655	2.363	1.833	1.179	1.001
	EWMA	Mean	220.788	29.779	7.384	3.712	2.370	1.380	1.073	1.001	1.000
		Median	290.172	42.363	10.032	4.889	3.048	1.694	1.211	1.005	1.000
		Midrange	140.017	32.355	9.020	4.536	2.882	1.625	1.181	1.008	1.000
		HL	260.660	33.133	7.900	3.916	2.495	1.428	1.092	1.001	1.000
		Trimean	300.958	35.151	8.143	4.015	2.552	1.454	1.099	1.001	1.000
		Trimmed	275.763	32.019	7.607	3.828	2.439	1.418	1.097	1.001	1.000
	Mixed EWMA–CUSUM	Mean	292.322	36.840	17.435	12.534	10.119	7.592	6.227	4.767	3.189
		Median	475.142	46.897	19.939	14.062	11.254	8.392	6.860	5.181	3.728
		Midrange	280.581	39.496	18.468	13.209	10.638	7.966	6.535	4.962	3.530
		HL	415.371	39.541	18.153	13.036	10.514	7.888	6.469	4.923	3.457
		Trimean	414.258	38.303	17.427	12.506	10.080	7.549	6.184	4.710	3.212
		Trimmed	851.272	47.058	19.921	14.111	11.326	8.465	6.920	5.240	3.747
10	CUSUM	Mean	212.117	16.842	5.714	3.499	2.588	1.852	1.330	1.007	1.000
		Median	291.308	23.585	6.827	3.973	2.869	1.977	1.495	1.009	1.000
		Midrange	78.245	23.243	8.796	5.241	3.767	2.511	2.003	1.305	1.085
		HL	260.353	19.338	6.027	3.622	2.655	1.886	1.366	1.004	1.000
		Trimean	273.492	20.145	6.030	3.592	2.625	1.857	1.325	1.003	1.000
		Trimmed	277.029	19.642	5.975	3.578	2.617	1.840	1.340	1.003	1.000
	EWMA	Mean	224.377	14.445	4.113	2.194	1.489	1.047	1.001	1.000	1.000
		Median	301.759	19.693	5.317	2.738	1.804	1.138	1.009	1.000	1.000
		Midrange	90.298	20.550	6.582	3.454	2.254	1.337	1.071	1.004	1.000
		HL	274.427	15.485	4.302	2.276	1.532	1.055	1.001	1.000	1.000
		Trimean	283.828	15.900	4.399	2.325	1.561	1.061	1.001	1.000	1.000
		Trimmed	288.092	15.799	4.397	2.321	1.558	1.064	1.002	1.000	1.000
	Mixed EWMA–CUSUM	Mean	294.426	24.424	13.203	9.765	7.960	6.023	4.976	3.918	2.938
		Median	503.884	28.285	14.623	10.678	8.659	6.519	5.333	4.030	2.987
		Midrange	216.461	28.891	15.075	11.009	8.949	6.732	5.533	4.168	2.982
		HL	428.580	25.382	13.677	10.105	8.235	6.223	5.104	3.966	2.957
		Trimean	455.428	25.495	13.578	10.007	8.145	6.152	5.050	3.938	2.926
		Trimmed	898.986	29.327	15.184	11.128	9.045	6.815	5.604	4.186	3.000

4.2. Performance comparison of the classical and robust control structures

In section we have described the six different location estimators which will be used in this study. In this subsection we will evaluate these estimators with the design structures of the CUSUM, EWMA and mixed EWMA–CUSUM control charts. The ARL_0 and ARL_1 based comparisons of the charts under the different environments discussed in section are given in Tables III–VI.

Keeping in mind that ARL_0 is a measure for robustness and ARL_1 is a measure of efficiency, the following points cover the findings of Tables III–VI:

1. Normal environment: As we have explained earlier we have taken $ARL_0=370$. It may be concluded from Table III that, if there is no shift, indeed the ARL_0 is around 370. Clearly, the ARL_1 performance of the EWMA control chart is best among the three control charts. This makes the EWMA control chart dominant over the CUSUM and mixed EMWA–CUSUM charts as the *ARLs* for EWMA are smaller for larger values of δ (cf. Table III). Under the uncontaminated normal environment, as it was expected, the sample mean performs best with the design structures of the CUSUM, EWMA and mixed EWMA–CUSUM charts as compared to all other estimators used (cf. Table III). For small shifts in the process, i.e. $\delta=0.25$, the EWMA chart with the trimmed mean estimator performs well, followed by the HL estimator with the EWMA structure. For small sample sizes, the mixed EWMA–CUSUM chart is slightly better than the CUSUM chart (cf. Table III).

Table V. ***ARL*** values of CUSUM, EWMA and mixed EWMA–CUSUM charts with $m=50$ at $ARL_0=370$ under localized variance contaminated normal environment

			ARL(δ)								
n	Chart	Estimator	0	0.25	0.5	0.75	1	1.5	2	3	5
5	CUSUM	Mean	186.939	34.273	9.005	5.036	3.558	2.325	1.862	1.132	1.000
		Median	183.256	45.501	11.868	6.240	4.257	2.687	2.065	1.377	1.001
		Midrange	187.388	41.651	10.939	5.962	4.142	2.640	2.046	1.358	1.001
		HL	189.957	35.935	9.582	5.313	3.726	2.418	1.921	1.195	1.000
		Trimean	183.341	36.725	9.414	5.132	3.581	2.324	1.842	1.125	1.000
		Trimmed	166.674	34.709	9.161	5.043	3.519	2.294	1.786	1.142	1.000
	EWMA	Mean	206.444	30.071	7.439	3.711	2.371	1.378	1.071	1.002	1.000
		Median	205.883	39.461	9.932	4.881	3.059	1.698	1.218	1.008	1.000
		Midrange	207.653	36.833	9.228	4.555	2.870	1.607	1.169	1.005	1.000
		HL	207.087	31.781	7.884	3.939	2.495	1.433	1.091	1.002	1.000
		Trimean	235.093	34.154	8.168	4.030	2.546	1.450	1.097	1.002	1.000
		Trimmed	187.948	30.133	7.663	3.829	2.448	1.422	1.098	1.003	1.000
	Mixed EWMA–CUSUM	Mean	290.984	37.045	17.397	12.540	10.114	7.588	6.225	4.767	3.187
		Median	291.370	43.480	19.349	13.695	10.957	8.169	6.678	5.057	3.625
		Midrange	293.450	41.961	19.189	13.645	10.978	8.202	6.716	5.084	3.688
		HL	291.633	38.454	17.954	12.887	10.397	7.795	6.384	4.885	3.345
		Trimean	288.958	37.402	17.301	12.398	9.984	7.467	6.125	4.665	3.110
		Trimmed	289.030	38.031	17.549	12.551	10.119	7.574	6.206	4.708	3.290
10	CUSUM	Mean	191.302	17.000	5.517	3.374	2.485	1.800	1.240	1.001	1.000
		Median	184.090	22.743	6.673	3.899	2.814	1.946	1.459	1.007	1.000
		Midrange	191.757	28.383	8.326	4.799	3.416	2.258	1.824	1.090	1.000
		HL	190.179	17.692	5.744	3.482	2.561	1.846	1.299	1.002	1.000
		Trimean	185.838	18.452	5.763	3.470	2.541	1.815	1.271	1.002	1.000
		Trimmed	179.613	18.033	5.718	3.461	2.543	1.800	1.289	1.002	1.000
	EWMA	Mean	209.400	14.597	4.121	2.190	1.478	1.044	1.003	1.000	1.000
		Median	209.208	19.153	5.304	2.747	1.812	1.141	1.012	1.000	1.000
		Midrange	209.875	24.620	6.724	3.420	2.211	1.304	1.047	1.001	1.000
		HL	207.123	15.368	4.333	2.285	1.536	1.057	1.003	1.000	1.000
		Trimean	206.977	15.600	4.407	2.338	1.570	1.063	1.004	1.000	1.000
		Trimmed	200.900	15.547	4.398	2.327	1.567	1.069	1.005	1.000	1.000
	Mixed EWMA–CUSUM	Mean	293.249	24.346	13.207	9.768	7.963	6.024	4.978	3.920	2.940
		Median	291.893	27.122	14.117	10.338	8.401	6.320	5.173	3.990	2.991
		Midrange	294.094	31.787	16.180	11.753	9.517	7.153	5.881	4.422	3.013
		HL	290.902	24.995	13.473	9.949	8.114	6.133	5.041	3.959	2.976
		Trimean	291.016	24.873	13.303	9.807	7.983	6.035	4.976	3.910	2.930
		Trimmed	289.129	25.112	13.446	9.913	8.071	6.098	5.012	3.872	2.860

2. Diffuse symmetric variance contaminated normal environment: We see from Table IV, that the ARL_0 of the mixed EWMA–CUSUM chart with the HL estimator is not much affected by the presence of a diffuse symmetric variance contamination in the distribution. On the other hand, the in-control *ARLs* for the other charts are disturbed, and the gap between the pre-fixed ARL_0 (370) and the attained ARL_0 is mostly significant. In most cases the number of false alarms has increased when there is a diffuse symmetric variance contamination present in the data but $\delta=0$. The effect of this kind of variance contamination on the design structure of these charts is obvious. For example, the in-control ARL_0 of the CUSUM, EWMA and mixed EWMA–CUSUM charts with the sample mean as location estimator decreases respectively by 44.58 %, 40.53 % and 20.93 % for $n=5$ and approximately the same decrease can be seen for $n=10$. This shows that the mixed EWMA–CUSUM is more robust to diffuse symmetric variance contaminations as compared to the CUSUM and EWMA charts. In the other words, the mixed EWMA–CUSUM reacts when there is a shift in the location parameter and does not react unnecessarily in the presence of variance contamination when we assume that this variance contamination is a part of the in-control process.
3. Localized variance contaminated normal environment: Table V reads that the in-control *ARL* s of the mixed EWMA–CUSUM chart is less affected as compared to the in-control *ARL* s of the CUSUM and EWMA charts. The influence of the estimator is limited for all charts when there are localized variance contaminations (the trimmed mean has the worst performance in the in-control situation). The CUSUM chart is producing many false alarms when the process is in-control. According to the ARL_1 determinant, the EWMA chart is very effective in finding out-of-control situations.

Table VI. ***ARL*** values of CUSUM, EWMA and mixed EWMA–CUSUM charts with $m=50$ at $ARL_0=370$ under diffuse asymmetric variance contaminated normal environment

			ARL(δ)								
n	Chart	Estimator	0	0.25	0.5	0.75	1	1.5	2	3	5
5	CUSUM	Mean	24.471	11.733	6.191	4.139	3.115	2.165	1.769	1.107	1.000
		Median	221.194	33.230	10.131	5.735	4.023	2.602	2.019	1.348	1.000
		Midrange	15.477	10.459	6.496	4.498	3.441	2.386	1.918	1.307	1.000
		HL	110.292	20.474	7.749	4.733	3.444	2.321	1.865	1.165	1.000
		Trimean	37.723	14.067	6.713	4.291	3.177	2.179	1.757	1.102	1.000
		Trimmed	130.813	22.258	7.672	4.577	3.298	2.212	1.741	1.124	1.000
	EWMA	Mean	24.765	10.197	4.986	3.014	2.092	1.307	1.053	1.000	1.000
		Median	229.483	27.581	8.391	4.370	2.840	1.627	1.188	1.003	1.000
		Midrange	15.167	9.325	5.382	3.457	2.433	1.494	1.133	1.001	1.000
		HL	116.648	17.650	6.174	3.407	2.274	1.367	1.072	1.000	1.000
		Trimean	42.583	12.694	5.629	3.292	2.246	1.373	1.076	1.000	1.000
		Trimmed	146.130	19.240	6.270	3.403	2.268	1.369	1.076	1.000	1.000
	Mixed	Mean	49.725	20.309	13.646	10.719	9.006	7.025	5.872	4.564	3.119
	EWMA–	Median	267.080	33.471	17.685	12.970	10.563	7.973	6.562	5.005	3.580
	CUSUM	Midrange	29.304	17.770	13.015	10.578	9.038	7.182	6.070	4.747	3.486
		HL	170.946	27.386	15.885	11.972	9.858	7.526	6.222	4.807	3.286
		Trimean	75.458	22.265	14.178	10.942	9.126	7.043	5.854	4.505	3.070
		Trimmed	214.485	28.489	15.887	11.840	9.691	7.368	6.081	4.645	3.257
10	CUSUM	Mean	17.351	7.027	3.978	2.789	2.196	1.653	1.178	1.000	1.000
		Median	223.303	16.272	5.874	3.617	2.676	1.897	1.412	1.002	1.000
		Midrange	7.919	5.992	4.329	3.293	2.654	1.968	1.652	1.061	1.000
		HL	134.052	11.551	4.903	3.182	2.412	1.778	1.247	1.000	1.000
		Trimean	143.537	12.270	4.970	3.179	2.401	1.750	1.223	1.000	1.000
		Trimmed	180.328	12.645	5.024	3.210	2.416	1.745	1.244	1.000	1.000
	EWMA	Mean	16.346	5.747	2.879	1.831	1.346	1.026	1.000	1.000	1.000
		Median	227.155	13.632	4.553	2.503	1.696	1.114	1.006	1.000	1.000
		Midrange	7.409	5.220	3.509	2.432	1.812	1.214	1.030	1.000	1.000
		HL	136.495	9.693	3.541	2.035	1.433	1.038	1.001	1.000	1.000
		Trimean	147.900	10.218	3.677	2.094	1.462	1.044	1.001	1.000	1.000
		Trimmed	186.448	10.652	3.731	2.103	1.470	1.049	1.001	1.000	1.000
	Mixed	Mean	33.412	14.895	10.454	8.352	7.084	5.561	4.687	3.710	2.795
	EWMA–	Median	247.442	22.434	13.132	9.880	8.123	6.186	5.097	3.974	2.989
	CUSUM	Midrange	15.376	11.530	9.346	7.974	7.025	5.764	4.961	3.926	2.872
		HL	168.006	19.863	12.254	9.365	7.761	5.962	4.961	3.923	2.959
		Trimean	183.772	19.981	12.175	9.270	7.660	5.871	4.895	3.860	2.900
		Trimmed	213.733	20.632	12.428	9.431	7.783	5.950	4.930	3.835	2.838

4. Diffuse asymmetric variance contaminated normal environment: From Table VI, we see that the in-control *ARL* s of the CUSUM, EWMA and the mixed EWMA–CUSUM charts are highly affected by the presence of such type of contaminations. The best in-control behavior is obtained with the median as estimator. Again the mixed EWMA–CUSUM chart has the best in-control *ARLs* and the EWMA chart is more efficient in detecting shifts as the corresponding ARL_1 s are the smallest (cf. Table VI).

To obtain a more global view of the run length distribution, along with the *ARL*, different indicators like the standard deviation of the run length (SDRL) and percentiles (denoted by P_i, $i=5, 25, 50, 75, 95$) of the run lengths of the in-control process are reported in Table VII. These measures help studying the short and long run behavior of the run length distribution. For instance, the 5% percentiles of the run length distribution of the CUSUM, EWMA and mixed EWMA–CUSUM charts are on average about 20, 14, and 40 observations for all estimators used. (cf. Table VII).

To get more insight into the out-of-control run length distribution, Figure 1 presents the run length distribution curves of all the charts considering $n=10$, $m=50$, $k=0.5$ and $\lambda=0.13$ with $\delta=0.25$ for normal environment. We only use three estimators: sample mean, sample median and HL estimator. In Figure 1, C, E and M represent, respectively, CUSUM, EWMA and mixed EWMA–CUSUM charts. The curves give the cumulative probability of detecting an out-of-control situation. A higher curve shows the superiority of a chart in terms of its quick detection of shifts in the process parameter. It can be observed from Figure 1 that EWMA charts based

Table VII. Characteristics of in-control run length distribution under uncontaminated normal environment for $n = 10$, $m = 50$, $k = 0.5$ and $\lambda = 0.13$ at $ARL_0 = 370$

Chart	Estimator	SDRL	Min	P_5	P_{25}	P_{50}	P_{75}	P_{95}	Max
CUSUM	Mean	441.06	3	21	87	220	482.25	1223.05	6516
	Median	425.46	2	21	95	234	499	1179	4488
	Midrange	414.8	2	24	100	238	497	1171	5086
	HL	443.5	3	20	89	220	484	1259.05	5047
	Trimean	424.06	2	21	90	222	487	1173	6226
	Trimmed	628.88	2	16	70	176	431	1322.2	16 482
EWMA	Mean	463.69	1	14	82	209	478	1265.1	7765
	Median	451.07	1	14	87	222	492	1245	5868
	Midrange	427.63	1	15	91	231	497	1221	6477
	HL	445.82	1	15	83.75	221	489.25	1250.05	4362
	Trimean	454.56	1	14	81	209	478	1235.1	7772
	Trimmed	592.98	1	11	65	175	435	1342	12 857
Mixed EWMA–CUSUM	Mean	467.16	15	40	93	199.5	442	1270	5056
	Median	443.42	14	40	99	216	466	1217.05	6810
	Midrange	398.36	16	45	110	230	478	1168	5197
	HL	453.89	16	41	96	204	468	1249.2	5500
	Trimean	442	13	40	95	207.5	455.25	1214	4772
	Trimmed	497.97	14	39	89	194	445	1262.15	6974

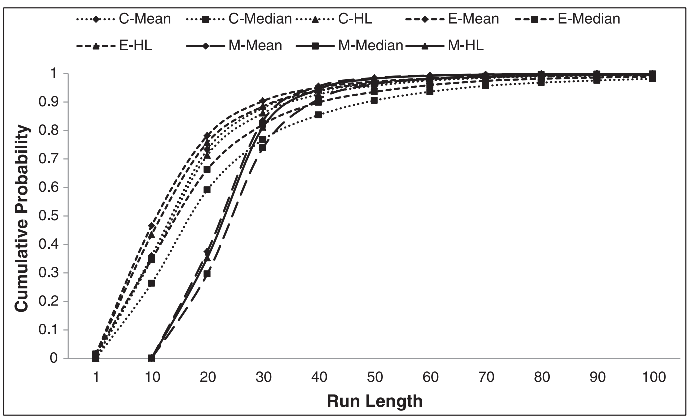

Figure 1. Run length curves for memory charts under uncontaminated normal environment when $n = 10$, $m = 50$, $k = 0.5$, $\lambda = 0.13$ and $\delta = 0.25$ at $ARL_0 = 370$

on all estimators have higher probabilities for small run lengths to detect the shift than those of other memory charts under normality. For detecting a shift of magnitude $\delta = 0.25$ at a run length equal to 50, the mixed EMWA–CUSUM has larger probabilities as compared to the EWMA and CUSUM charts.

5. Summary and conclusion

Control charts are widely used in monitoring and controlling variations present in the process location and dispersion. Commonly applied control charts are the memory-less (Shewhart-type) charts for targeting the large shifts and memory (EWMA and CUSUM) charts for aiming the smaller shifts. A combination of the EWMA and CUSUM control charts is applied to enhance the performance of the charts even further. The current study presents a comparison of the CUSUM, EWMA and mixed EWMA–CUSUM control charts based on different estimators. Different parent environments (normal and contaminated normal) are used to evaluate the performance of these charts in terms of their *ARLs* and different percentiles of the *RL* distribution. The comparisons showed that there is no single control chart or estimator which behaves well in all environments. Under normality the EWMA control chart based on the sample mean is the best, although the differences with the other charts and estimators are insignificant (especially for small shifts). When there are localized or diffuse symmetric variances contaminations the mixed EWMA–CUSUM control chart is quite robust against these variance contaminations. Overall the best performance is obtained by the EWMA control chart based on the median estimator.

Acknowledgement

The co-authors, Muhammad Riaz and Nasir Abbas, are indebted to the King Fahd University of Petroleum and Minerals Dhahran Saudi Arabia for providing excellent research facilities.

References

1. Page ES. Continuous inspection schemes. *Biometrika* 1954; **41**(1–2):100–115.
2. Roberts SW. Control chart tests based on geometric moving averages. *Technometrics* 1959; **1**(3):239–250.
3. Lucas JM. Combined Shewhart–CUSUM quality control schemes. *Journal of Quality Technology* 1982; **14**(2):51–59.
4. Lucas JM, Saccucci MS. Exponentially weighted moving average control schemes: properties and enhancements. *Technometrics* 1990; **32**(1):1–12.
5. Steiner SH. EWMA control charts with time varying control limits and fast initial response. *Journal of Quality Technology* 1999; **31**(1):75–86.
6. Capizzi G, Masarotto G. An adaptive exponentially weighted moving average control chart. *Technometrics* 2003; **45**(3):199–207.
7. Zhao Y, Tsung F, Wang Z. Dual CUSUM control schemes for detecting a range of mean shifts. *IIE Transactions* 2005; **37**(11):1047–1057.
8. Colosimo BM, Godio F, Palmieri L. Comparative studies of control charts for torque data in automotive component assembling. *International Journal of Technology Management* 2007; **37**(1):72–85.
9. Castagliola P, Celano G, Fichera S, Giuffrida F. A variable sampling interval S2–EWMA control chart for monitoring the process variance. *International Journal of Technology Management* 2007; **37**(1):125–246.
10. Machado MAG, Costa AFB. The double sampling and the EWMA charts based on the sample variances. *International Journal of Production Economics* 2008; **114**(1):134–148.
11. Riaz M, Abbas N, Does RJMM. Improving the performance of CUSUM charts. *Quality and Reliability Engineering International* 2011; **27**(4):415–424.
12. Abbas N, Riaz M, Does RJMM. Enhancing the performance of EWMA charts. *Quality and Reliability Engineering International* 2011; **27**(6):821–833.
13. Abbas N, Riaz M, Does RJMM. Mixed exponentially weighted moving average-cumulative sum charts for process monitoring. *Quality and Reliability Engineering International* 2013; **29**(3):355–365.
14. Woodall WH, Montgomery DC. Research issues and ideas in statistical process control. *Journal of Quality Technology* 1999; **31**:376–386.
15. Jensen WA, Jones-Farmer LA, Champ CW, Woodall WH. Effects of parameter estimation on control charts properties: a literature review. *Journal of Quality Technology* 2006; **38**(4):349–364.
16. Schoonhoven M, Nazir HZ, Riaz M, Does RJMM. Robust location estimators for the Xbar chart. *Journal of Quality Technology* 2011; **34**(4):363–379.
17. Nazir HZ, Riaz M, Does RJMM, Abbas N. Robust CUSUM control charting. *Quality Engineering* 2013; **25**(3):211–224.
18. Dixon WJ, Massey FJ. Introduction to statistical analysis (3rd edn). McGraw-Hill: New York, 1969.
19. Ferrell EB. Control charts using midranges and medians. *Industrial Quality Control* 1953; **9**(5):30–34.
20. Hettmansperger TP, McKean JW. Robust nonparametric statistical methods. Wiley: New York, 1998.
21. Wang T, Li Y, Cui H. On weighted randomly trimmed means. *Journal of Systems Science and Complexity* 2007; **20**(1):47–65.
22. Song MS, Chung HY, Bae W. Subset selection procedures based on some robust estimators. *Journal of the Korean Statistical Society* 1982; **11**:109–117.
23. Caperaa P, Rivest LP. On the variance of trimmed mean. *Statistics and Probability Letters* 1995; **22**(1):79–85.
24. Khattree R, Rao CR. *Statistics in industry—(handbook of statistics; v. 22)* (1st edn). Elsevier: The Netherlands, 2003.
25. Crowder SV. Design of exponentially weighted moving average schemes. *Journal of Quality Technology* 1989; **21**(3):155–162.

Authors' biographies

Hafiz Zafar Nazir obtained his MSc in Statistics from the Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan, in 2006, and MPhil in Statistics from the Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan, in 2008. He obtained his PhD in statistics from the Institute for Business and Industrial Statistics, University of Amsterdam, the Netherlands, in September, 2014. He is serving as a lecturer in the Department of Statistics, University of Sargodha, Pakistan, from September 2009 to till date. His current research interests include statistical process control, nonparametric techniques, and robust methods.

Nasir Abbas did his PhD in Statistics from the Institute for Business and Industrial Statistics, University of Amsterdam, the Netherlands, in 2008. Currently, he holds the position of assistant professor in the Department of Mathematics and Statistics, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia. His current research interests include statistical process control.

Muhammad Riaz obtained his PhD in statistics from the Institute for Business and Industrial Statistics, University of Amsterdam, the Netherlands, in 2008. He holds the position of associate professor in the Department of Mathematics and Statistics, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia. His current research interests include statistical process control, nonparametric techniques, and experimental design.

Ronald J.M.M. Does is a professor of Industrial Statistics at the University of Amsterdam; director of the Institute for Business and Industrial Statistics, which operates as an independent consultancy firm within the University of Amsterdam; head of the Department of Operations Management; and director of the Institute of Executive Programmes at the Amsterdam Business School. He is a fellow of the ASQ and ASA and Academician of the International Academy for Quality. His current research activities include the design of control charts for nonstandard situations, health care engineering, and operations management methods.