# Control charts for location based on different sampling schemes

Mehmood, R.; Riaz, M.; Does, R.J.M.M.

**DOI**
10.1080/02664763.2012.740624

**Publication date**
2013

**Document Version**
Final published version

**Published in**
Journal of Applied Statistics


Journal of Applied Statistics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/cjas20

# Control charts for location based on different sampling schemes

Rashid Mehmood [a] , Muhammad Riaz [b c] & Ronald J.M.M. Does [d]

[a] Department of Commerce, Islamia University Bahawalpur Campus, Rahim Yar Khan, Pakistan

[b] Department of Mathematics and Statistics, King Fahad University of Petroleum and Minerals, Dhahran, 31261, Saudi Arabia

[c] Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan

[d] Department of Quantitative Economics, IBIS UvA, University of Amsterdam, Plantage Muidergracht 12, 1018 TV, Amsterdam, The Netherlands
Version of record first published: 05 Dec 2012.

To cite this article: Rashid Mehmood , Muhammad Riaz & Ronald J.M.M. Does (2013): Control charts for location based on different sampling schemes, Journal of Applied Statistics, 40:3, 483-494

To link to this article: http://dx.doi.org/10.1080/02664763.2012.740624

*Journal of Applied Statistics*, 2013
Vol. 40, No. 3, 483–494, http://dx.doi.org/10.1080/02664763.2012.740624





# Control charts for location based on different sampling schemes

Rashid Mehmood[a], Muhammad Riaz[b,c*] and Ronald J.M.M. Does[d]

[a]*Department of Commerce, Islamia University Bahawalpur Campus, Rahim Yar Khan, Pakistan;* [b]*Department of Mathematics and Statistics, King Fahad University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia;* [c]*Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan;* [d]*Department of Quantitative Economics, IBIS UvA, University of Amsterdam, Plantage Muidergracht 12, 1018 TV Amsterdam, The Netherlands*

(*Received 26 February 2012; final version received 13 October 2012*)

Control charts are the most important statistical process control tool for monitoring variations in a process. A number of articles are available in the literature for the $\bar{X}$ control chart based on simple random sampling, ranked set sampling, median-ranked set sampling (MRSS), extreme-ranked set sampling, double-ranked set sampling, double median-ranked set sampling and median double-ranked set sampling. In this study, we highlight some limitations of the existing ranked set charting structures. Besides, we propose different runs rules-based control charting structures under a variety of sampling strategies. We evaluate the performance of the control charting structures using power curves as a performance criterion. We observe that the proposed merger of varying runs rules schemes with different sampling strategies improve significantly the detection ability of location control charting structures. More specifically, the MRSS performs the best under both single- and double-ranked set strategies with varying runs rules schemes. We also include a real-life example to explain the proposal and highlight its significance for practical data sets.

**Keywords:** control charts; power; runs rules; sampling strategies; statistical process control

## 1. Introduction

In 1920s, Shewhart (cf. [28]) introduced the idea of control charts to monitor variations in the process parameters, including location and spread. Following the pioneering work of Shewhart, different Shewhart-type control charts were developed for the process mean and dispersion using the simple random sampling (SRS) approach. McIntyre [16] suggested ranked set sampling (RSS) to estimate the population mean, and Takahasi and Wakimoto [30] derived the necessary mathematical theory in this regard. RSS has wide application in different areas, including agriculture, medicine, sociology and ecology for different purposes, cf. [1,14,19,21,23,24,27] and the references therein. Dell and Clutter [8] started the discussion in which the ranking may not be perfect, that is, there are errors in ranking the units with respect to the variable of interest.

*Corresponding author. Email: riaz76qau@yahoo.com

Shewhart-type control charts with different sampling strategies generally perform well for large shifts in the process parameters. In case of SRS, a number of authors have defined sensitizing rules and runs rules schemes for different control charts to improve the detecting ability for small and moderate shifts. These rules are described and evaluated in the Western Electric Statistical Quality Control Handbook [31], and in the articles of Khoo [11], Khoo and Ariffin [12], Acosta-Mejia [3], Koutras *et al.* [13], Antzoulakos and Rakitzis [6], Acosta-Mejia and Pignatiello [4] and Riaz *et al.* [25], among others.

Although the implementation of these runs rules improves the performance of control charts, it creates problems such as inflated false alarm rates, biasedness and non-monotonicity (cf. [25]). Several authors have addressed these issues in different ways such as simultaneous application of more rules or adjusting the structure of the control limits. Riaz *et al.* [25] redefined the rules for the sake of providing each rule an independent identity and at the same time handling the aforementioned issues.

In this article, we implement different runs rules schemes to the design structure of control charts for location based on different sampling strategies. It includes SRS, RSS, median-ranked set sampling (MRSS), extreme-ranked set sampling (ERSS), percentile-ranked set sampling (PRSS), double-ranked set sampling (DRSS), double median-ranked set sampling (DMRSS), median double-ranked set sampling (MDRSS), double percentile-ranked set sampling (DPRSS) and percentile double-ranked set sampling (PDRSS). Furthermore, we carry out extensive comparisons among these strategies under the runs rules schemes.

The organization of this article is as follows: In Section 2, we provide the conceptual framework of the different sampling strategies. In Section 3, we work out the design structures of the proposed runs rules-based control charts under different sampling strategies for monitoring the location. Then, in Section 4, we evaluate and compare the proposed control charts using power curves as performance measure. In Section 5, we provide a numerical example for illustrative purposes. Finally, Section 6 provides the summary, and the main findings, and gives a direction for future research.

## 2. Different sampling strategies

In practical situations, we may have some extra information about the study variable readily available or it can easily be gathered. To make sampling mechanisms more effective, McIntyre [16] proposed the idea of RSS, which uses extra information by ranking the units, and several authors suggested to use control charts based on different ranked set strategies, such as RSS, MRSS and ERSS (see, for example, [21] and the references therein). The ranked set methods are mainly classified into two categories, namely single- and double-ranked set sampling strategies. Later, we discuss both categories (along with their different variants).

The RSS procedure can be summarized as (cf. [21]):

> Select $n$ random sample sets, each of size $n$ units from the population, and rank the units with respect to variable of interest. Then an actual measurement is taken from the unit with the smallest rank from the first set. From the second set, an actual measurement is taken from the unit with the second smallest rank, and the procedure is continued until the unit with the largest rank is chosen for actual measurement from the $n$th set. This way we obtain a sample of $n$ measured units, one from each set. The cycle may be repeated until $nr$ units have been measured.

This results in simple RSS.

If the median-order statistics is chosen from each set in the above procedure, we get the MRSS. Another strategy is the PRSS where any particular percentile (e.g. the 30th percentile) is selected from each set. Also, ERSS is a strategy in which the smallest observation is picked from the first half and the largest observation is picked from the next half. A double application of these strategies results in DRSS strategies such as DRSS, DMRSS and DPRSS. Moreover, we get MDRSS if we choose RSS at the first stage and then choose the median-ordered observations at the second stage. Similarly, after choosing an RSS in the first stage and if we pick any particular percentile at the second stage, it results in PDRSS. For more details, the interested reader may study Muttlak and Saleh [22], Muttlak [20], [1,2,5,21] and the references therein.

Sometimes, the variable of interest $X$ is difficult to rank (cf. [8]), in which case we rank this variable with respect to another variable, the so-called concomitant variable ($Y$). The procedure is described as: select $n$ bivariate random samples each of size n. Rank the variable of interest ($X$) with respect to the concomitant variable ($Y$). For an actual measurement from the first set, the unit with the smallest $Y$ is associated, from the second set select the unit corresponding with the second smallest $Y$, continue this procedure until we obtain the unit corresponding with the largest $Y$. This results in the so-called imperfect ranked set sampling (IRSS). The idea of IRSS may also be worked out for the other single- and double-ranked set strategies on similar lines.

Let $X_{(i:n)j}$ be $i$th order statistic of the quality characteristic of interest $X$ for the $i$th SRS set of size $n$ in the $j$th cycle. Then, an unbiased estimator of the population mean (cf. [30]) using ranked set sample data based on the $j$th cycle is defined by: $\bar{X}_{R,j} = 1/n \sum_{i=1}^{n} X_{(i:n)j}$; $j = 1, 2, \ldots, r$ and the standard deviation is given by $\sigma_{R,j} = \sqrt{1/n^2 \sum_{i=1}^{n} \sigma^2_{(i,n)}}$, where $\sigma^2_{(i:n)} = \mathrm{E}((X_{(i:n)} - E(X_{(i:n)}))^2$ represents the variance of the $i$th order statistic from a sample of size $n$.

Assuming that $X$ and $Y$ have a bivariate normal distribution then the regression of $X$ on $Y$ is defined as [29] $X = \mu_X + \rho(\sigma_X/\sigma_Y)(Y - \mu_Y) + \varepsilon$, where $\varepsilon$ is the random error component having mean zero and variance $\sigma_X^2(1-\rho^2)$, where $\rho$ is the correlation between $X$ and $Y$ and $\mu_X, \mu_Y, \sigma_X, \sigma_Y$ are the means and standard deviations of the variables $X$ and $Y$, respectively, and $Y$ and $\varepsilon$ are independently distributed. Let $(X_{[i,n],j}, Y_{(i:n),j})$ denote the pair of $i$th smallest value of $Y$ associated with the corresponding value of $X$ obtained from the $i$th set in the $j$th cycle, respectively. Hence, we have: $X_{[i,n],j} = \mu_X + \rho(\sigma_X/\sigma_Y)(Y_{(i,n),j} - \mu_Y) + \varepsilon_{ij}$. Then, the unbiased estimator of the population mean for the variable of interest $X$ with concomitant variable $Y$ can be written as: $\bar{X}_{\mathrm{IR},j} = 1/n \sum_{i=1}^{n} X_{[i:n],j}$ $j = 1, 2, \ldots, r$ and the standard deviation is given by: $\sigma_{\bar{X}_{\mathrm{IR},j}} = \sqrt{\sigma_X^2/n\,[(1-\rho^2) + \rho^2/n \sum_{i=1}^{n} \sigma^2_{z(i:n)}]}$, where $\sigma^2_{z(i:n)}$ represents the variance of the $i$th order statistic from a sample of size $n$ from the standard normal distribution.

The unbiased estimators of the population mean for the other single- and double-ranked set strategies (both for the perfect and imperfect ranking strategies) may be defined on the similar lines by using the respective ranked data of each sampling strategy accordingly, as discussed above for RSS and IRSS. For more relevant details, the interested readers may see Garcia and Cebrian [10], Salazar and Sinha [26], Muttlak and Saleh [22], Muttlak [20], [1,7,21,28,29] and the references therein.

## 3. Proposed control charting schemes

To monitor the location parameter of the variable under study, the general structure of the location control chart is given as: $\hat{\mu} \pm K\sigma_{\hat{\mu}}$, where$\hat{\mu}$ is a location estimator such as the usual $\bar{X}$ based on SRS or any of the RSS estimators $\hat{\mu}$ as given in Section 2, $\sigma_{\hat{\mu}}$ is the corresponding standard deviation of the respective estimator (e.g. $\sigma_X/\sqrt{n}$ for the usual $\bar{X}$) as provided in Section 2. $K$ is the control limit coefficient, which fixes the false alarm rate. The first RSS location control chart structure was given by $\mu \pm 3\sqrt{\sigma_X^2/n\ [(1-\rho^2) + \rho^2/n \sum_{i=1}^{n} \sigma^2_{z(i:n)}]}$. This form was extensively studied by Muttlak and Al-Sabah [21]. They used the factor 3 as control limit coefficient and implemented only one out-of-control rule, that is, if one sample $\bar{X}_R$ falls outside the control limits, then we have an out-of-control signal (we call this a 1 out of 1 rule which will symbolically be written as 1/1). In fact, the coefficient 3 is meant for the usual SRS-based $\bar{X}$ control chart, so using the same coefficient with the design structure of the RSS-based $\bar{X}_R$ control chart results in three limitations:

(1) The false alarm rate cannot be fixed at the desired level;
(2) The design structure cannot be used for varying choices of $n$ and $\rho$ purposefully (as alarm rates are disturbed);
(3) The decision rule is restricted to only 1/1.

The same approach is followed in the literature for the other sample schemes described in Section 2, and hence these results have the same limitations.

We propose in this article extended design structures for the location control chart based on the different ranked set strategies from Section 2, namely RSS, MRSS, ERSS, PRSS, DRSS, DMRSS, DPRSS, MDRSS, PDRSS and we will use different runs rules schemes. The resulting control charts will take care of the above-mentioned three limitations as well. For convenience, we will use the notation $T$ to refer to different sampling strategies such as RSS, MRSS and ERSS. Before providing the mentioned design structures, we list down the set of runs rules schemes, which will be implemented with the design structure of the suggested control charts. These runs rules are given as in Riaz *et al.* [25], who applied these rules with the SRS-based control charts:

- An out-of-control signal is received if at least *k–m* points of k consecutive points (where $0 \le m \le k-1$) fall outside the control (signalling) limits $(h_L, h_U)$ of the sampling distribution of the control charting statistic $\bar{X}_T$. Symbolically, we write these rules as: $k - m/k$.

Here,

$$h_L = \mu_X + A_{(T,n,\rho,m,k,p/2)}\sigma_X, \quad h_U = \mu_X + A_{(T,n,\rho,m,k,1-p/2)}\sigma_X, \tag{1}$$

where the control limits constant $A_{(T,n,\rho,m,k,p/2)}$ depends on the sampling strategy $T$ (which may be any of the strategies as listed in Section 2), the sample size $n$, and the correlation $\rho$ between $X$ and $Y$; $k$ are the total consecutive points to be considered, *k–m* are the decision observations used in a given rule and $p$ is the probability of a single point falling outside the respective signalling limits depending on *k–m* and $k$. The value of $p$ for a given runs rule depends on $m$, $k$ and $\alpha$ (the pre-specified false alarm rate) and may be obtained by solving the following equation:

$$\alpha = \sum_{k-m\le k} \frac{k!}{(k-m)!m!} p^{(k-m)}(1-p)^m, \text{ where } 0 \le m \le k-1. \tag{2}$$

The values of $p$ for different rules investigated in this study are listed in Table 1 for $\alpha = 0.0027$. Similar values may be obtained for any choice of $\alpha$ by solving Equation (2).

The control structure given in Equation (1) represents different control charting structures, namely the $\bar{X}_R$, $\bar{X}_M$, $\bar{X}_E$, $\bar{X}_P$ $\bar{X}_{\mathrm{DR}}$, $\bar{X}_{\mathrm{DM}}$, $\bar{X}_{\mathrm{DP}}$, $\bar{X}_{\mathrm{MD}}$ and $\bar{X}_{\mathrm{PD}}$ control charts based on their respective

Table 1. Values of $p$ for different sensitizing rules.

| Rule | 1/1 | 2/2 | 1/2 | 3/3 | 2/3 | 1/3 | 4/4 | 3/4 |
|---|---|---|---|---|---|---|---|---|
| $p$ | 0.0027 | 0.0014 | 0.1392 | 0.1392 | 0.0303 | 0.0009 | 0.008 | 0.0898 |
| Rule | 2/4 | 5/5 | 4/5 | 3/5 | 6/6 | 5/6 | 4/6 | 7/7 |
| $p$ | 0.0215 | 0.1577 | 0.0669 | 0.0669 | 0.3732 | 0.2231 | 0.1219 | 0.4296 |
| Rule | 6/7 | 5/7 | 8/8 | 7/8 | 6/8 | 9/9 | 8/9 | 7/9 |
| $p$ | 0.2826 | 0.4774 | 0.3554 | 0.3554 | 0.2304 | 0.5183 | 0.3821 | 0.2788 |

Table 2. Unbiasing constants for different ranked set-based charts.

| $\rho$ | 0.25 | | | 0.50 | | | 0.75 | | | 1 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | 3 | 5 | 7 | 3 | 5 | 7 | 3 | 5 | 7 | 3 | 5 | 7 |
| RSS | 1.699 | 2.33 | 2.717 | 1.747 | 2.377 | 2.754 | 1.815 | 2.447 | 2.816 | 1.939 | 2.564 | 2.956 |
| MRSS | 1.675 | 2.261 | 2.637 | 1.57 | 2.114 | 2.412 | 1.406 | 1.8 | 2.007 | 1.144 | 1.296 | 1.368 |
| ERSS | 1.725 | 2.4 | 2.326 | 1.812 | 2.546 | 2.343 | 1.95 | 2.802 | 2.335 | 1.939 | 3.082 | 3.769 |
| DRSS | 1.704 | 2.327 | 2.73 | 1.76 | 2.368 | 2.76 | 1.864 | 2.466 | 2.848 | 2.026 | 2.671 | 3.03 |
| DMRSS | 1.646 | 2.272 | 2.624 | 1.515 | 2.043 | 2.368 | 1.263 | 1.624 | 1.834 | 0.757 | 0.665 | 0.568 |
| MDRSS | 1.656 | 2.259 | 2.616 | 1.537 | 2.052 | 2.367 | 1.303 | 1.655 | 1.879 | 0.89 | 0.86 | 0.786 |
| PRSS | 1.675 | 2.267 | 2.637 | 1.57 | 2.115 | 2.455 | 1.406 | 1.824 | 2.064 | 1.144 | 1.296 | 1.368 |
| DPRSS | 1.646 | 2.245 | 2.638 | 1.515 | 2.051 | 2.37 | 1.263 | 1.626 | 1.87 | 0.758 | 0.736 | 0.721 |
| PDRSS | 1.659 | 2.253 | 2.645 | 1.527 | 2.062 | 2.385 | 1.306 | 1.689 | 1.929 | 0.894 | 0.932 | 0.953 |

sampling strategies RSS, MRSS, ERSS, PRSS, DRSS, DMRSS, DPRSS, MDRSS, PDRSS, for monitoring the location parameter $\mu_X$. These structures are capable of addressing the limitations (i)–(iii) mentioned above and also take care of the issues of biasedness and non-monotonicity as studied by Riaz *et al.* [25]. Moreover, these control structures help implementing each rule independently and hence avoiding simultaneous application (cf. [9]) of many rules, which results in a simpler design structure.

For the case of unknown parameters, Equation (1) may be written as

$$h_L = \bar{\bar{X}}_T + A_{(T,n,\rho,m,k,p/2)} \frac{\bar{R}_T}{d_{2(T)}}, \quad h_U = \bar{\bar{X}}_T + A_{(T,n,\rho,m,k,(1-p/2))} \frac{\bar{R}_T}{d_{2(T)}},$$

where $\bar{\bar{X}}_T$ represents the mean of the corresponding sample means, $\bar{R}_T$ represents mean of the sample ranges and $d_{2(T)}$ represents the unbiasing constant for a given strategy $T$. For some selective choices, the values of $d_{2(T)}$ are provided in Table 2.

The control limit coefficient $A_{(T,n,\rho,m,k,p/2)}$ for the $\bar{X}_T$ control charts, for a pre-specified $\alpha$, may be obtained for given design parameters $n$, $\rho$, $m$ and $k$. For some selective choices of these quantities, we have obtained the values of $A_{(n,\rho,m,k,1-p/2)}$ for different preferences of $T$ as given in Section 2. The results are tabulated in the form of Table 3 for RSS by fixing $\alpha = 0.0027$. Similar tables may be constructed for other sampling strategies at different choices of the design parameters. We have also generated more tables for MRSS, ERSS, PRSS, DRSS, DMRSS, DPRSS, MDRSS, PDRSS; and for the sake of brevity, these results are reported in Tables 4–11 at the weblink http://www.ibisuva.nl/en/Research in the form of "Additional material". The results of Tables 2–11 are obtained through extensive Monte Carlo simulations ($10^5$–$10^6$ runs depending on the stability of estimate and the reduction of error in a given situation). To conduct a Monte Carlo experiment, we have generated random samples using a specific sampling strategy $T$, for a given size, from the standard bivariate normal distribution (without loss of generality) with varying values of $\rho$. Bivariate normality may be seen in different industrial applications (cf. [15]). The choices of $\rho$ result in different ranking mechanisms, such as $\rho = 0$ leads to SRS, $\rho = 1$ leads to perfect ranking, and $\rho$ between 0 and 1 results in imperfect ranking (cf. [1]).

For a given sample, we have computed the values of the random variable $\bar{X}_T$ for each combination of $\rho$ and $n$ and for a given sample scheme $T$. By repeating these steps $10^5$–$10^6$ times using different combinations of $\rho$ and $n$, we obtain values for $A_{(n,\rho,m,k,p/2)}$ according to the pre-specified false alarm rate $\alpha$. Note that for a desired $\alpha$, the value of $p$ is obtained by solving Equation (2). The choices of runs rules provided in Tables 3–11 are motivated by Riaz *et al.* [25]. The results for other choices may easily be derived on similar lines.

Table 3. Control limits of RSS for different run rules at $\alpha = 0.0027$.

| $\rho$ | $n$ | Control limits | Rule 1/1 | 2/3 | 2/4 | 9/9 | 8/9 | 7/9 |
|---|---|---|---|---|---|---|---|---|
| 0 | 3 | $A_{(n,\rho,m,k,p/2)}$ | −1.73277 | −1.25056 | −1.3269 | −0.37328 | −0.50423 | −0.6246 |
| | | $A_{(n,\rho,m,k,1-p/2)}$ | 1.73407 | 1.249735 | 1.326254 | 0.372247 | 0.504329 | 0.624287 |
| | 5 | $A_{(n,\rho,m,k,p/2)}$ | −1.3411 | −0.96894 | −1.02879 | −0.28976 | −0.39162 | −0.48504 |
| | | $A_{(n,\rho,m,k,1-p/2)}$ | 1.343263 | 0.969543 | 1.027768 | 0.288665 | 0.390497 | 0.484182 |
| | 7 | $A_{(n,\rho,m,k,p/2)}$ | −1.13494 | −0.81872 | −0.86916 | −0.24428 | −0.33017 | −0.40931 |
| | | $A_{(n,\rho,m,k,1-p/2)}$ | 1.136441 | 0.819057 | 0.869706 | 0.244302 | 0.330421 | 0.409436 |
| 0.25 | 3 | $A_{(n,\rho,m,k,p/2)}$ | −1.69895 | −1.23277 | −1.30686 | −0.3668 | −0.49687 | −0.61471 |
| | | $A_{(n,\rho,m,k,1-p/2)}$ | 1.705628 | 1.232535 | 1.30811 | 0.368274 | 0.497586 | 0.616615 |
| | 5 | $A_{(n,\rho,m,k,p/2)}$ | −1.31354 | −0.94859 | −1.00689 | −0.28389 | −0.38394 | −0.47511 |
| | | $A_{(n,\rho,m,k,1-p/2)}$ | 1.314215 | 0.948729 | 1.007537 | 0.283188 | 0.38267 | 0.473988 |
| | 7 | $A_{(n,\rho,m,k,p/2)}$ | −1.10681 | −0.79944 | −0.84791 | −0.2386 | −0.32293 | −0.39996 |
| | | $A_{(n,\rho,m,k,1-p/2)}$ | 1.104576 | 0.79944 | 0.848399 | 0.238572 | 0.322971 | 0.400112 |
| 0.5 | 3 | $A_{(n,\rho,m,k,p/2)}$ | −1.6293 | −1.17502 | −1.24719 | −0.34958 | −0.47318 | −0.58731 |
| | | $A_{(n,\rho,m,k,1-p/2)}$ | −1.6293 | −1.17502 | −1.24719 | −0.34958 | −0.47318 | −0.58731 |
| | 5 | $A_{(n,\rho,m,k,p/2)}$ | −1.22811 | −0.88804 | −0.94243 | −0.26528 | −0.35892 | −0.4444 |
| | | $A_{(n,\rho,m,k,1-p/2)}$ | 1.230368 | 0.888501 | 0.942625 | 0.264525 | 0.357863 | 0.443675 |
| | 7 | $A_{(n,\rho,m,k,p/2)}$ | −1.02776 | −0.73976 | −0.78576 | −0.22074 | −0.29881 | −0.37019 |
| | | $A_{(n,\rho,m,k,1-p/2)}$ | 1.024953 | 0.741068 | 0.786638 | 0.221262 | 0.299232 | 0.370535 |
| 0.75 | 3 | $A_{(n,\rho,m,k,p/2)}$ | −1.48647 | −1.07045 | −1.13683 | −0.31964 | −0.43198 | −0.53534 |
| | | $A_{(n,\rho,m,k,1-p/2)}$ | 1.480112 | 1.071179 | 1.135862 | 0.31865 | 0.43137 | 0.534283 |
| | 5 | $A_{(n,\rho,m,k,p/2)}$ | −1.07211 | −0.77578 | −0.82294 | −0.23124 | −0.31293 | −0.38754 |
| | | $A_{(n,\rho,m,k,1-p/2)}$ | 1.071079 | 0.774892 | 0.821796 | 0.231132 | 0.31294 | 0.387268 |
| | 7 | $A_{(n,\rho,m,k,p/2)}$ | −0.87244 | −0.63144 | −0.66984 | −0.18825 | −0.25462 | −0.31519 |
| | | $A_{(n,\rho,m,k,1-p/2)}$ | 0.878487 | 0.63145 | 0.670181 | 0.188902 | 0.25559 | 0.316106 |
| 1.00 | 3 | $A_{(n,\rho,m,k,p/2)}$ | −1.26359 | −0.9057 | −0.96181 | −0.26877 | −0.36405 | −0.45115 |
| | | $A_{(n,\rho,m,k,1-p/2)}$ | 1.261075 | 0.905653 | 0.960996 | 0.268683 | 0.36392 | 0.451047 |
| | 5 | $A_{(n,\rho,m,k,p/2)}$ | −0.81238 | −0.58488 | −0.62095 | −0.17333 | −0.23457 | −0.29076 |
| | | $A_{(n,\rho,m,k,1-p/2)}$ | 0.813623 | 0.583784 | 0.620062 | 0.172694 | 0.23381 | 0.289998 |
| | 7 | $A_{(n,\rho,m,k,p/2)}$ | −0.60286 | −0.43275 | −0.45924 | −0.12837 | −0.174 | −0.21554 |
| | | $A_{(n,\rho,m,k,1-p/2)}$ | 0.601519 | 0.431769 | 0.45881 | 0.128687 | 0.173935 | 0.215439 |

## 4. Performance evaluation and comparisons

This section evaluates and compares the performance of our different sampling strategies-based proposals in the form of $\bar{X}_R$, $\bar{X}_M$, $\bar{X}_E$, $\bar{X}_P$, $\bar{X}_{\mathrm{DR}}$, $\bar{X}_{\mathrm{DM}}$, $\bar{X}_{\mathrm{DP}}$, $\bar{X}_{\mathrm{MD}}$ and $\bar{X}_{\mathrm{PD}}$ control charts (depending on the choice of $T$) by implementing different runs rules schemes particularly those studied by Riaz *et al.* [25]. Power curves are used as performance criterion following Montgomery [18] and Riaz *et al.* [25]. For power computations in this article, we have considered shifts in terms of $\delta\sigma$, which means that the shifted location parameter say $\mu'$ is defined as: $\mu' = \mu + \delta\sigma$) and limits are set using the two-sided probability limits given in Equation (1), for specified $\alpha$,$n$, $\rho$, $m$ and $k$ using a particular choice of $T$. Here, $\delta = 0$ means no shift in $\mu$ and $\delta \neq 0$ means that the process $\mu$ has been changed. The power evaluations for different values of $n$, $\rho$ and different choices of $\alpha$ are obtained for all $\bar{X}_T$ control charts with varying options of $T$, $m$ and $k$. Particularly, we have considered the schemes with $k = 1, 2, \ldots, 9$ and $m = 0, 1, 2$ such that the set of rules "at least $k$–$m$ points of $k$ consecutive points (where $0 \leq m \leq k - 1$)" may be classified into three categories as those where $k$–$m = 0, 1, 2$. To be more specific, $k$–$m = 2$ includes 1/3, 2/4, 3/5, 4/6, 5/7, 6/8, 7/9; $k$–$m = 1$ includes 1/2, 2/3, 3/4, 4/5, 5/6, 6/7, 7/8, 8/9 and $k$–$m = 0$ includes 1/1, 2/2, 3/3, 4/4, 5/5, 6/6, 7/7, 8/8, 9/9. Of course, other choices of $k$ and $m$ may be opted but those considered here are able to cover the most practical and commonly used situations.

Assuming bivariate normality of $(X_i, Y_i)$ the power evaluations are carried out in the following way: for fixed values of $\alpha$, we have chosen the value of the control limit coefficient $A_{(T,n,\rho,m,k,p/2)}$ of the control structure for the $\bar{X}_T$ control chart for different choices of the design parameters $n$, $\rho$, $m$ and $k$ with some appropriately worked out value of $p$. Then, proportions of cases falling outside $(h_L, h_U)$ are calculated using their corresponding control charting statistics $\bar{X}_T$ (or $\bar{X}$ which is in fact a special case of $\bar{X}_R$ when $\rho = 0$) for varying values of $\delta$. We have accomplished this power study by designing a code in R language and by running extensive Monte Carlo simulations ($10^5$–$10^6$ depending on stability of the output).

For some selective values of the design parameters $n$, $\rho$, $m$ and $k$ at $\alpha = 0.0027$ and for varying values of $\delta$, power curves are provided in Figures 1–5 for different choices of $T$. We have also produced some more figures for MRSS, ERSS, PRSS, DRSS, DMRSS, DPRSS, MDRSS, PDRSS, but for limiting the number of pages, the extended figures can be found at the weblink http://www.ibisuva.nl/en/Research in the form of "Additional material". Note that for PRSS, we have considered 30th and 40th percentiles in our study. We have presented extreme choices of $m$ and $k$ for runs rules schemes in these figures and tables because the performance of the other rules may easily be covered as in between cases, as was already noted by Riaz *et al.* [25].

For the proposed run rules-based charting structures under varying rank set strategies, we have observed that: these are flexible to accommodate any value of the design parameters $n$, $\rho$, $m$ and



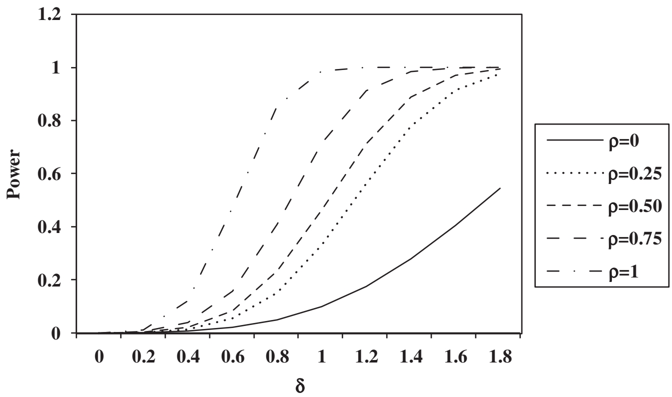


Figure 1. Power curves of different control charts with varying $\rho, n, r$, and $m$ at $\alpha = 0.0027$.

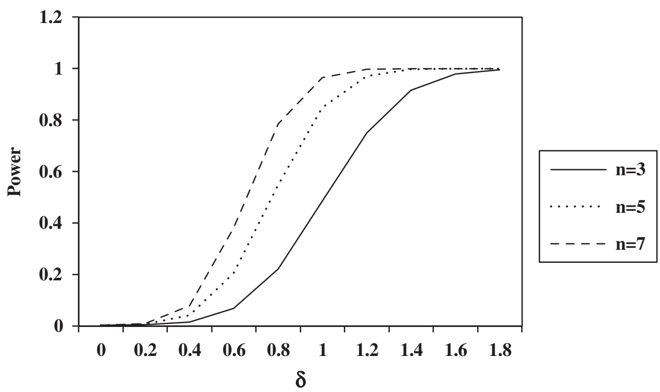


Figure 2. Power curves of different control charts using varying $n$ with different rules at $\alpha = 0.0027$.

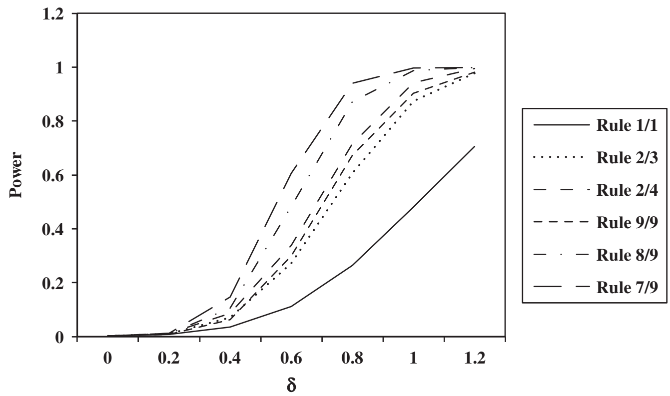


Figure 3. Power curves of different control charts with varying runs rules at $\alpha = 0.0027$.

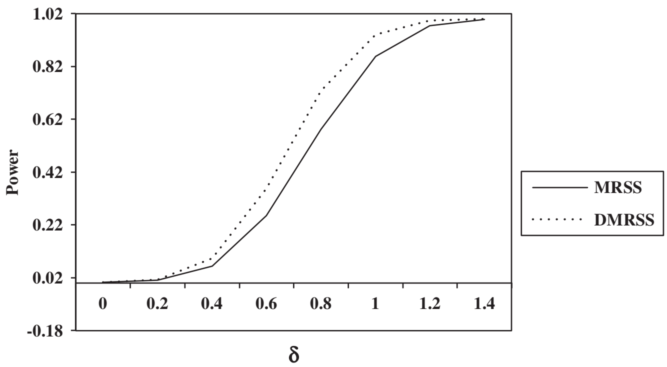


Figure 4. Power curves of double- versus single-ranked set strategies at $\alpha = 0.0027$.

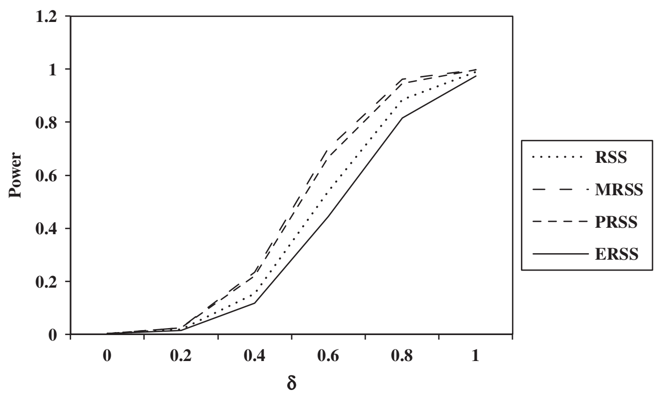


Figure 5. Power curves of different control charts with varying ranked set strategies using different $\rho, n, m$ and $k$ at $\alpha = 0.0027, n = 5$, rule 2/3 and $\rho = 1$.

$k$ at any desired pre-specified value of $\alpha$ (cf. Tables 2–11), which is not the case with simple RSS/DRSS and the usual $\bar{X}$ charts. They avoid inflation of false alarm rates for any choice of $n$ and $\rho$, and also for any runs rule choice (cf. Tables 3–11), which is a serious issue in runs rules application with any control charting structures such as RSS/DRSS/SRS-based location charts. Their performances remain superior as compared with those of the simple RSS/DRSS/SRS-based design structures, the usual $\bar{X}$ chart and the runs rules-based design structure of the $\bar{X}$ chart given by Riaz *et al.* [25], for varying choices of the design parameters $n$, $\rho$, $m$ and $k$ at any value of $\alpha$ (cf. Figures 1–5). Their performances keep improving with an increase in the values of $\rho$ (cf. Figure 1), $n$(cf. Figure 2), $k$(cf. Figure 3) and $k - m$ (cf. Figure 3). They avoid complicated design structures of the control chart, which occur due to simultaneous application of different runs rules. As all our implemented runs rules have independent identities (cf. [25]), there is no need to apply more rules simultaneously. Hence, the control charting structure remains simple and their power ability keep improving with an increase in the values of $\delta$ and $\alpha$, so there is no issue of biasedness and non-monotonicity as already indicated by Riaz *et al.* [25].

Moreover, the design structures of the simple RSS-based charts (cf. [21]) and the simple DRSS-based charts (cf. [1]) become special cases of the proposed structures given in Equation (1) when $m = k = 1$ and the choice of the control limit coefficient $A = 3$. The design structure of the usual $\bar{X}$ control chart also becomes a special case of the proposed structure given in Equation (1) when $m = k = 1$ and $\rho = 0$ and the design structure of the runs rules-based $\bar{X}$ chart given by Riaz *et al.* [25] becomes a special case of the proposed structure (1) when $\rho = 0$.

We have also examined that the double-ranked set strategies are better as compared with single-ranked set strategies in general (cf. Figure 4). In the class of single-ranked set strategies, the superiority order is MRSS, PRSS, ERSS and RSS, while in double-ranked set group, the dominance order is DMRSS, DPRSS, MDRSS, PDRSS and DRSS with varying runs rules schemes. It advocates that the runs rules implementation with different variants of RSS and DRSS really boost the performance of these control charting structures in general (cf. Figure 3).

In short, the design structure of the proposed $\bar{X}_T$ charts are generalized versions of the usual Shewhart $\bar{X}$ chart (cf. [28]), the RSS-based design structure of the $\bar{X}$ chart (cf. [21]), the DRSS-based design structure of the $\bar{X}$ chart (cf. [1]) and the runs rules-based design structure of the $\bar{X}$ chart (cf. [25]) .

## 5. A real life example

To illustrate the application of ranked set control charts and to highlight their importance by implementing a variety of runs rules schemes, we provide here a numerical example. In this study, we consider an agricultural scenario where the yield depends on many factors such as soil, water, environment and treatment. Among these, soil is an important factor which provides different nutrients, such as nitrogen, phosphorous, potassium, zinc, boron, iron, copper, which constitute the total soluble salt in soil. An optimum level of soluble salt is essential for good production. The presence/absence of soluble salt in soil can be assessed easily by visual inspection (concomitant variable). To monitor the soluble salt of different plots through a control chart, random samples are collected using MRSS because of its good performance as shown in Section 4.

The procedure of collecting the median ranked set sample is as follows: For 30 villages select 25 plots from each village using SRS and then divide them randomly into 5 sets of size 5 each. By visual inspection, assign the rank to each set with respect to their total soluble salt. For an actual measurement of total soluble salt, select the plot with rank 3 and measure its electric conductivity (EC) using a portable device (EC meter in ms/cm). Repeat this process 30 times (cycle). Hence, 30 median ranked set samples of size 5 each are gathered.

The objective of the study is to check whether the plots are suitable for cultivation and if in case of an effected plot (out-of-control) a chemist can recommend a remedy. In the example, the variable of interest is EC in soil and the ranking variable is the visual inspection of soluble salt in soil.

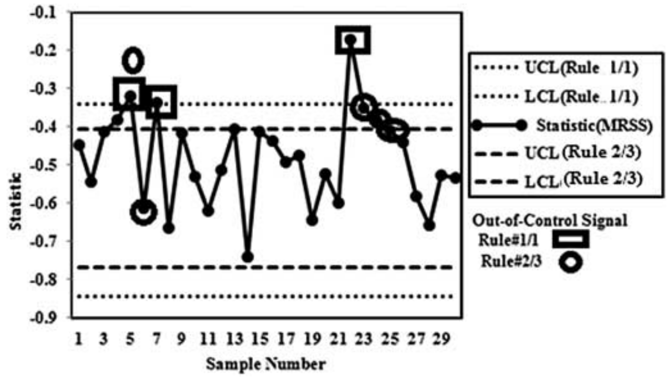


Figure 6. MRSS control chart with runs rules on EC data.

It is to be mentioned that the said data set was giving a closer fit to a lognormal distribution. As in our study, we assume normality of the process so for that purpose we performed a log transformation. We tested this data set using the MRSS-based control chart with the implementation of 1/1 and 2/3 runs rule schemes as an illustration. We have set two-sided control limits, namely $h_L$ and $h_U$ at $\alpha = 0.0027$ and two runs rules as:

$$A_{(n,\rho,m,k,p/2)} = -0.7182 \text{ and } A_{(n,\rho,m,k,1-p/2)} = 0.7150 \text{ for the 1/1 rule;}$$
$$A_{(n,\rho,m,k,p/2)} = -0.5189 \text{ and } A_{(n,\rho,m,k,1-p/2)} = 0.5187 \text{ for the 2/3 rule.}$$

The values of the control charting statistic $\bar{X}_M$ are calculated for all the 30 ranked set samples and are plotted against their respective limits $h_L$ and $h_U$ depending on the runs rule to be used. The resulting control chart for MRSS is presented in Figure 6, in which the symbols LCL and UCL are used to indicate lower and upper control limits, respectively. The transformed data set is plotted in the form of a figure (cf. Figure 7), provided on the weblink http://www.ibisuva.nl/en/Research as “Additional material” to save some space in the article. The original data set is also provided in the form of Table 12, on the weblink http://www.ibisuva.nl/en/Research.

The out-of-control signals on this chart are indicated by rectangular boxes for the 1/1 rule and oval shapes for the 2/3 rule. It is evident from Figure 6 that there are three out-of-control signals given using 1/1 rule and five signals using 2/3 rule. This indicates the significance of runs rules-based structures for these charts. It is to be noted that the 2/3 rule gives more signals as compared with the 1/1 rule, which is in accordance with the findings of Section 4.

## 6. Summary and conclusions

Control charts are important statistical tools to monitor process characteristics for stability, where the charting statistics are displayed on graphs by plotting these statistics versus the sampling order. A number of control charts have been introduced based on different sampling technique such as SRS and RSS, but their design structures partially cover the choices of different design parameters such as $n$, $\rho$, $m$, $k$ and any desired value of $\alpha$. The proposals of this study serve as a generalized version, which can accommodate a variety of choices for the said design parameters. Indeed, the design structure of the proposed charts are more saturated versions, which cover the features of the usual SRS along with single- and double-ranked set strategies (e.g. the usual Shewhart $\bar{X}$ chart [28], single RSS-based design structures [21], double RSS-based design structures [1] and runs rules-based design structure of $\bar{X}$ chart [25]. In general, double-ranked set strategies perform better as compared with single ranked set strategies and within single-ranked set strategies MRSS is the best choice followed by PRSS, ERSS and RSS, whereas in double-ranked set class DMRSS outperform all the others followed by DPRSS, MDRSS, PDRSS and DRSS.

The proposed control structures take care of different issues (such as inflated false alarm rate, biasedness, non-monotonicity and complicated design structure) and provide an improved efficiency for every type of shifts as compared with the existing counterparts.

A detailed investigation on dispersion charts with varying RSS schemes supported by different runs rules is a natural follow-up of this article. It is hard to find literature on dispersion charts with ranked set strategies, so the idea of location control charting with ranked set strategies implemented with runs rules schemes needs to be studied in detail with dispersion charts as well (cf. [17]).



## Acknowledgements

The authors are greatly thankful to the anonymous reviewers and the editor for their encouragement and valuable suggestions to further improve the initial version of the article. The author Muhammad Riaz is indebted to the King Fahd University of Petroleum and Minerals, Saudi Arabia for providing excellent research facilities under IN111059. The authors are also thankful to Soil and Water Testing Laboratory R.Y.K, Pakistan for their co-operation about EC example data.

## References

[1] M.R. Abujiya and H.A. Muttlak, *Quality control chart for the mean using double rank set sampling,* J. Appl. Stat. 31 (2004), pp. 1185–1201.
[2] M.R. Abujiya and H.A. Muttlak, *Some extensions to double ranked set sampling,* J. Appl. Statist. Sci. 15(3) (2007), pp. 365–379.
[3] C.A. Acosta-Mejia, *Two sets of runs rules for the X chart,* Qual. Eng. 19 (2007), pp. 129–136.
[4] C.A. Acosta-Mejia and J.J. Pignatiello, *ARL-designs of D charts with k-of-k runs rules*, Comm. Statist. Simul. Comput. 38 (2009), pp. 1625–1639.
[5] M.F. Al-Saleh and M. Al-Kadiri, *Double rank set sampling*, Statist. Probab. Lett. 48 (2000), pp. 205–212.
[6] D.L. Antzoulakos and A.C. Rakitzis, *The modified r out of m control chart*, Comm. Statist. Simul. Comput. 37 (2008), pp. 396—408
[7] Z. Chen, Z. Bai, and B.K. Sinha, *Rank Set Sampling,* Springer, New York, 2004.
[8] T.R. Dell and J.L. Clutter, *Ranked set sampling theory with order statistics background*, Biometrics 28 (1972), pp. 545–553.
[9] R.J.M.M. Does and B.F. Schriever, *Variables control chart limits and tests for special causes*, Statist. Neerlandica 46 (1992), pp. 229–245.
[10] M.R. Garcia and A.A. Cebrian, *Repeated substitution method: The ratio estimator for the population variance*, Metrika 43 (1996), pp. 101–105.
[11] M.B.C. Khoo, *Design of runs rules schemes,* Qual. Eng. 16 (2004), pp. 27–34.
[12] M.B.C. Khoo and K.N. Ariffin, *Two improved runs rules for the Shewhart $\bar{X}$ chart,* Qual. Eng. 18 (2006), 173–178.
[13] M.V. Koutras, S. Bersimis, and P.E. Maravelakis, *Statistical process control using Shewhart control chart with supplementary runs rules,* Methodol. Comput. Appl. Probab. 9 (2007), pp. 207–224.
[14] P.H. Kvam and F.J. Samaniego, *Nonparametric maximum likelihood estimation based on rank set sampling*, J. Amer. Statist. Assoc. 89 (1994), pp. 526–537.
[15] R.L. Mason and J.C. Young, *Multivariate Statistical Process Control with Industrial Applications*, ASA SIAM, Philadelphia, PA, 2002.
[16] G.A. McIntyre, *A method for unbiased selective sampling using ranked sets*, Aust. J. Agric. Res. 3 (1952), pp. 385–390.
[17] R. Mehmood, M. Riaz, and R.J.M.M. Does, *Dispersion control charting based on different sampling schemes*, Manuscript in preparation, (2012).
[18] D.C. Montgomery, *Introduction to Statistical Quality Control,* 6th ed., Wiley, New York, 2009.
[19] H.A. Muttlak, *Median ranked set sampling*, J. Appl. Statist. Sci. 6(4) (1997), pp. 245–255.
[20] H.A. Muttlak, *Investigating the use of quartile ranked set samples for estimating the population mean*, Appl. Math. Comput. 150 (2004), pp. 543–554.
[21] H.A. Muttlak and W.S. Al-Sabah, *Statistical quality control based on ranked set sampling*, J. Appl. Stat. 30 (2003), pp. 1055–1078.

[22] H.A. Muttlak and M.F. Saleh, *Recent developments on ranked set sampling*, Pak. J. Statist. 16 (2000), pp. 269–290.

[23] O. Ozturk, *Parametric estimation of location and scale parameters in ranked set sampling*, J. Statist. Plann. Inference 141(4) (2011), pp. 1616–1822.

[24] O. Ozturk and N. Balakrishnan, *An exact control-versus-treatment comparison test based on ranked set samples,* Biometrics 65(4) (2009), pp. 1213–1222.

[25] M. Riaz, R. Mehmood, and R.J.M.M. Does, *On the performance of different control charting rules*, Qual. Reliab. Eng. Int. 27(8) (2011), pp. 1059–1067.

[26] R.D. Salazar and A.K. Sinha, *Control chart* $\bar{X}$ *based on ranked set sampling*, Comun. Tecica No. 1-97-09(PE/CIMAT), 1997.

[27] H. Samawi, W. Abu-Dayyeh, and S. Ahmad, *Extreme rank set sampling*, Biometric. J. 30 (1996), pp. 577–586.

[28] W. Shewhart, *Economic Control of Quality Manufactured Product*, Van Nostrand, New York, 1931; reprinted by the American Society for Quality Control in 1980, Milwaukee, WI.

[29] S.L. Stokes, *Ranked set sampling with concomitant variables*, Comm. Statist. Theory Methods A6(12) (1977), pp. 1207–1211.

[30] K. Takahasi and K. Wakimoto, *On the unbiased estimates of the population mean based on the sample stratified by means of ordering,* Ann. Inst. Statist. Math. 20 (1968), pp. 1–31.

[31] Western Electric Company, *Statistical Quality Control Handbook,* Western Electric Company, Indianapolis, 1956.