SUPER QC-PACK: A NOVEL STATISTICAL SOFTWARE FOR ANALYSING QUALITY CONTROL DATA

Waheed Babatunde Yahya, Isaac Adeola Adeniyi

Department of Statistics, University of Ilorin, P.M.B. 1515, Ilorin, Nigeria

ABSTRACT: In this paper, we present a novel statistical package (Super QC-Pack) for analysing quality control data and constructing various quality control charts. The new Super QC-Pack (SQC) computes the control limits for various quality control charts and simultaneously provides the graphical outputs of the results in an easily understood manner. The SQC package is currently packed as a single setup which can be installed directly into any windows operating system. It was developed using the C-Sharp programming language under the Visual Studio .Net 2008 framework and the Zed graph graphics library. The SQC package is computationally efficient and provides accurate results like any of the existing statistical packages. The new SQC software is graphics-driven and very user-friendly that is suitable for users with limited background in statistics and statistical computing. The SQC software is not commercialised. It is freely available and can be downloaded and installed without restrictions at www.unilorin.edu.ng/sqcpack/yahya_adeniyi/webpage.htm

KEYWORDS: P-chart, C-chart, R-chart, S-chart, U-chart, X-Bar chart, process capability indices, C-sharp, Microsoft Visual Studio 2008, Zedgraph graphics library.

1. INTRODUCTION

In statistical process control (SPC), the stability of a production or service process is monitored by control charts. The SPC methods extend the use of descriptive statistics and the concept of hypothesis testing to monitor the quality of products and processes. The most versatile tools adopted in SPC to monitor the stability of a process are the control charts ([OS90, Mon05, DRT12]). A control chart is used to monitor, control, and improve the performance of manufacturing or service process over time by studying variation and its source.

Various characteristics in manufacturing or service processes that are examined through control charts can be divided into two groups of variables and attributes ([B+07])**.** The *control chart for variables* is used to monitor characteristics that can be measured on continuum scales such as height, weight, volume and so on. Example of this include the X-Bar chart which is used to monitor average process performance as well as the R and S charts both of which are designed to monitor process variation ([Mon05, Bau07]). There also exist individual X-charts, Moving Range chart, standardized X-Bar chart and Exponentially Weighted Moving-Average (EWMA) Control Chart ([Mon85, Mon91, LS90]) each of which is equally meant to monitor specific process characteristic.

The *control chart for attributes* is used to monitor qualitative characteristics in a production process like taste, colour, smell and the like, in which only the number of objects that possess or do not possess the require quality characteristic(s) of interest are observed. An attribute requires only a single decision, such as yes or no, good or bad, acceptable or unacceptable, defective or non-defective and so on. The most commonly adopted control charts under this category are the P-chart for proportion of defectives (non-conforming), the C-chart for number of defects and the U-chart for number of defects when the group size is greater than one.

Most often, it is of interest to evaluate the ability of a manufacturing or service process at meeting or exceeding the preset specifications set by the management of the enterprise. This can be simply accomplished by examining some of the *process capability indexes* (PCI) such as C_p, C_{pk}, C_{pm}, P_p and P_{pk} among others details of which have been provided elsewhere ([Mon85, Mon91, Mon05, KJ92, B+07]).

It is quite instructive to note that, the tasks involve in the construction of various control charts require the users' understanding of some of the basic statistical theories and their applications. Also, the large volume of data that emanate on daily basis from various manufacturing processes has made it necessary (if not mandatory) for the quality control officers (QCOs) of these manufacturing outfits to be highly skilful in the usage of any of the existing statistical software for data processing and analysis. These would enable them to monitor the stability of the production processes of their industries efficiently. Interestingly, a good number of the existing statistical packages like SPSS, R, MINITAB, SAS and the like assumed that their users already possess some basic requisite knowledge in statistical theories and their applications.

In many food processing and/or pharmaceutical companies in Nigeria for instance, the QCOs are

Chemists, Biochemists or Microbiologists who have limited knowledge of statistical methods and their applications. Unfortunately, these are the people that are often charged with the responsibilities of monitoring and taking records of the production outputs, processing and analysing the data that emanated from the process and reporting their findings to the management for necessary actions. However, due to lack of the needed requisite skills in statistics and statistical computing, it might be difficult for some of the QCOs of these companies to efficiently analyse the data that emanated from their production processes using some of the existing statistical software. It is the recognition of all these constraints and many more that motivated the development of the new statistical software presented in this work to cater for this category of users and statisticians alike who usually savour good graphical display of their results.

This paper therefore presents a novel Super QC-pack (SQC) software, version 1.0.0 for analysing quality control data. The new SQC software is a user-friendly statistical package that computes all the necessary statistics for the construction of quality control charts and provides the graphical outputs of the results computed in form of the intended control charts. One of the basic features of this new software lies in its flexibility and simplicity. The software can be easily used by people with very limited statistical background.

The new SQC is designed purposely to cater for all categories of users in manufacturing or service outfits, universities and colleges and the academia.

2. THE DEVELOPMENT OF THE SUPER QC-PACK SOFTWARE

The new Super QC-pack software version 1.0.0 was developed for windows platform machines using the C sharp .NET programming language, an object oriented language, and using the Visual Studio 2008 environment as the designer. The Zedgraph graphics library (www.sourceforge.net/projects/zedgraph/) was employed to make the graph controls easier to handle. The software has been deployed and packed with .NET windows installer that makes it very easy to install and run independent of the visual studio environment.

The introduction of Visual Studio .NET and the birth of the C-sharp programming language ([ET02]) has led to the development of programs with more powerful features and better graphics as evident in this work.

A 10.7MB setup file of the new SQC software can be downloaded and installed freely at www.unilorin.edu.ng/sqcpack/yahya_adeniyi/webpage.htm. Detail on system requirements for the proper installation of the software are provided in the next sub-section. However, the authors can be contacted for further information on the usage of the software.

2.1. *System Requirements and Installation*

The SQC version 1.0.0 can be installed and run on any 32 or 64 bits Windows XP, Windows VISTA, Windows 7 and Windows server 2003/2008 operating systems. A minimum of 256MB of RAM with at least 17MB hard disc space is required for optimal performance.

The installation is simply done by double clicking the installer (setup file). A windows installer would pop up, click next to proceed, select users ("Just Me" or "Everyone") and choose installation folder, click next to finalize your installation. After the installation, the "Shortcut to Super QC-Pack" icon will be deposited on the desktop and in the user's program menu. Click on this icon to launch the SQC software.

The installation of SQC package is automatically accompanied by a comprehensive but self-instructive manual, in portable document format (pdf) that contained detail instructions on how to run the various analyses using the software. Information on data preparation for analysis using the package is also provided in the manual. To view the pdf manual click on "All programs" and locate "Super QC-Pack". Finally, click on "Super QC-pack manual" to download or view the manual.

3. THE CONSTRUCTIONS OF BASIC CONTROL CHARTS

Control charts ([Dun86, Que93, Wil06]) are one of the most commonly used methods in statistical process control to monitor the stability of a manufacturing or service process. It is a visual display of the fluctuations of the measured quantities or attributes from a process that will enable the quality control officers to determine whether these variations fall within the specified process limits. Therefore, a process is either classified as being in statistical quality control or out of statistical quality control depending on the type of pattern display by the plotted process units around the two bounds called the *upper control limits* (UCLs) and *lower control limits* (LCLs).

Without loss of generality therefore, the construction of any control chart requires the computation of mean value of the process (center line), standard deviation and range of a set of process data collected when the process is under stable operation. All these statistics are used to determine the values of UCL and LCL of the manufacturing or service process. Data collected from subsequent processes can then

be plotted and compared with the calculated UCL and LCL to determine whether the new processes are in control or out of control.

The six basic control charts implemented by SQC pack are X-bar chart, R-chart and S-chart for variable measurements as well as P-chart, C-chart and U-chart for attributes measurements. Also, the three process capability indices that can be computed using SQC pack are C_p, C_{pk} and C_{pm}.

Each of the charts mentioned above was developed based on the underlying distribution of the type of data (process) it is meant to monitor. The X-Bar, R, and S charts were developed based on the assumption that the data are normally distributed. The P-chart is based on the assumption that the data are from binomial distribution, the C-chart and U-chart are based on the assumption that the data are from Poisson distribution. All these underlying distributions were employed as efficient tools to estimate the center line (CL), UCL and LCL for the construction of each control chart to develop the SQC software. All the computations are based on the procedures reported in Montgomery (2005) and the computation summary of each procedure is provided in the following subsection as a guide to users.

3.1 *Variable Control Charts*

The X-Bar Chart

An X-bar chart ([Mon04; Mon05]) is quality control tool to monitor the mean value of a process over time. For a given data set, the *center line* (CL), UCL and LCL are computed and the x-bar values in each group or batch is plotted. The upper and lower control limits define the range of inherent variation in the group means when the process is in control.

The *center line* (CL) is the grand mean of all the groups/batches given by μ which, if it is unknown a priori from a stable process can be estimated from the data by

$$\overline{\overline{X}} = \frac{1}{m}\sum_{j=1}^{m}\overline{X}_j \qquad (1)$$

where $\overline{X}_j$ is the mean of samples in group/batch *j*, *j* = 1, 2, ..., *m*. The *upper control limit* (UCL) of the process is determined by

$$\mu + Z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \qquad (2)$$

while its *lower control limit* (LCL) is determined by

$$\mu - Z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \qquad (3)$$

where σ is the standard deviation of the process under stable condition, n is the number of sample in each group and $Z_{1-\alpha/2}$ is the quantile of the standard normal distribution at significance level α.

Like the center line μ, the value of σ is not usually known and can only be estimated from the samples collected from the process. The following two statistics are available to estimate σ:

$$\hat{\sigma} = \frac{\overline{R}}{d_2} \qquad (4)$$

or

$$\hat{\sigma} = \frac{\overline{S}}{c_4} \qquad (5)$$

where $\overline{R}$ is the average of all the m groups ranges and $\overline{S}$ is the average of all the m groups standard deviations which are estimated by $\overline{R} = \frac{1}{m}\sum_{j=1}^{m} R_j$ and $\overline{S} = \frac{1}{m}\sum_{j=1}^{m} S_j$ respectively. Both R_j and S_j are the range and standard deviation for group j which are respectively determined by:

$$R_j = \max(X_j) - \min(X_j)$$

and

$$S_j = \frac{1}{\sqrt{n-1}}\sqrt{\sum_{i=1}^{n}\left(x_{ij} - \overline{x}_{ij}\right)^2}.$$

The denominators d_2 and c_4 in (4) and (5) are constants that depend on n (each group sample size). The estimate of d_2 can be obtained ([Mon05]) by

$$d_2 = \int_{-\infty}^{\infty}[1 - (1 - \Phi(x))^n - \Phi(x)^n]dx$$

where

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-(x^2/2)}dx$$

Also, c_4 is estimated by

$$c_4 = \sqrt{\left(\frac{2}{n-1}\right)\frac{\left(\frac{n-2}{2}\right)!}{\left(\frac{n-3}{2}\right)!}}$$

However, for n greater 25, c_4 are approximated by $(4n-3)/(4n-4)$.

The product $Z_{1-\alpha/2} * \sigma/\sqrt{n}$ in (2) and (3) which is controlled by the size of Type I error α defines the number of sigma used to compute the UCL and LCL. Thus, control charts with 1-sigma, 2-sigma and 3-sigma employed about 68.26%, 95.44% and 99.73% confidence coefficients at α values of approximately 0.31, 0.05 and 0.003 respectively.

For unequal group sample sizes n_j, the centre line (CL) is computed by

$$\overline{\overline{X}} = \frac{1}{\sum_{j=1}^{m} n_j} \sum_{j=1}^{m} \sum_{i=1}^{n} x_{ij} \qquad (6)$$

Here, individual UCL and LCL are computed for each group j by

$$\text{UCL} = \overline{\overline{X}} + Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n_j}} \qquad (7)$$

$$\text{LCL} = \overline{\overline{X}} - Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n_j}} \qquad (8)$$

The estimate of σ in (7) and (8) is determined by $\frac{1}{m}\sum_{j=1}^{m} \frac{R_j}{d_2(n_j)}$ or $\frac{1}{m}\sum_{j=1}^{m} \frac{S_j}{c_4(n_j)}$ depending on whether the range or standard deviation of samples in each group is used respectively. The values of $d_2(n_j)$ and $c_4(n_j)$ in the above quantities are determined by substituting n_j for n in their respective estimators as earlier stated.

The R-chart

The R-chart ([CW87]) is used to monitor the variation in production or service process. These charts will allow us to see any deviations from desired limits within the quality process and, in effect, allow the firm to make necessary adjustments to improve quality. The construction of R-chart requires the computations of centre line (CL) $\overline{R}$, and the upper and lower control limits given by

$$\text{UCL} = \overline{R}\left(1 + Z_{1-\frac{\alpha}{2}} \frac{d_3}{d_2}\right) \qquad (9)$$

$$\text{LCL} = \overline{R}\left(1 - Z_{1-\frac{\alpha}{2}} \frac{d_3}{d_2}\right) \qquad (10)$$

where $\overline{R}$ and d_2 are as defined and used in equation (4). The numerator d_3 in the expressions (9) and (10) is another constant that depends on the group sample size n and it is estimated ([Mon05]) by

$$d_3 = \left[2\int_{-\infty}^{\infty}\int_{-\infty}^{x}[1-\Phi(x)^n-(1-\Phi(y))^n+(\Phi(x)-\Phi(y))^n]dydx - d_2^2\right]^{\frac{1}{2}}$$

where

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-(x^2/2)}dx$$

and

$$\Phi(y) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} e^{-(y^2/2)}dy$$

For varying group sample sizes, the following estimators for the construction of the R-chart are employed:

$$\text{CL} = \overline{R} = \frac{1}{\sum_{j=1}^{m} n_j}\sum_{j=1}^{m} R_j \qquad (11)$$

The upper and lower control limits for samples in group j are given by

$$\text{UCL} = \overline{R} + Z_{1-\frac{\alpha}{2}}\sigma d_3(n_j) \qquad (12)$$

$$\text{LCL} = \overline{R} - Z_{1-\frac{\alpha}{2}}\sigma d_3(n_j) \qquad (13)$$

where σ is estimated by $\frac{1}{m}\sum_{j=1}^{m}\frac{R_j}{d_2(n_j)}$ while d_3 is as earlier stated but only evaluated at each group sample sizes $n_j, j = 1,\ldots,m$.

The S-Chart

Like R-chart, S-chart is also used to monitor the variation in production or service processes using the

standard deviations of the process. To construct the S-chart, the following statistics are required:

$$\text{CL} = \bar{S} = \frac{1}{m}\sum_{j=1}^{m} S_j \tag{14}$$

$$\text{UCL} = \bar{S}\left(1 + Z_{1-\frac{\alpha}{2}}\frac{\sqrt{1-c_4^2}}{c_4}\right) \tag{15}$$

$$\text{LCL} = \bar{S}\left(1 - Z_{1-\frac{\alpha}{2}}\frac{\sqrt{1-c_4^2}}{c_4}\right) \tag{16}$$

where S_j is the standard deviation of samples in group j and constant c_4 are as defined and used in equation (5).
For varying group sample sizes n_j, the estimators in (14), (15), and (16) become

$$\text{CL} = \bar{S} = \frac{1}{\sum_{j=1}^{m} n_j}\sum_{j=1}^{m} S_j \tag{17}$$

$$\text{UCL} = \bar{S} + Z_{1-\frac{\alpha}{2}}\sigma\sqrt{1-c_{4(n_j)}^2} \tag{18}$$

$$\text{LCL} = \bar{S} - Z_{1-\frac{\alpha}{2}}\sigma\sqrt{1-c_{4(n_j)}^2} \tag{19}$$

respectively where σ is estimated by

$$\frac{1}{m}\sum_{j=1}^{m}\frac{S_j}{c_4(n_j)}$$

and $c_4(n_j)$ is as earlier stated but evaluated at each group sample sizes n_j.

3.2 *Attribute Control Charts*

Control charts for attributes ([Woo97, RS97, Rya00]) are quality control tools used to monitor the quality characteristics that represent attributes of products. The most common of these are the P-chart, C-chart, np-chart, and U-chart.

The P-Chart

The P-chart is a type of attribute control charts that is most applicable to situations where the occurrence of defectives is not rare. To use this chart, the percentage of defectives may be expected to be more than 5% of the total number of units produced.
In the p-chart chart, the percentage of defectives (per batch) over time is plotted. However, the control limits are based on the binomial distribution (of proportions). The CL, UCL and LCL are estimated by the following:

$$\text{CL} = \bar{p} = \frac{1}{\sum_{j=1}^{m} n_j}\sum_{j=1}^{m} d_j \tag{20}$$

where d_j is the number of defectives in group/batch j, and n_j is the corresponding sample size in group/batch j.
The control limits are

$$\text{UCL} = \bar{p} + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{p}(1-\bar{p})}{n}} \tag{21}$$

$$\text{LCL} = \bar{p} - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{p}(1-\bar{p})}{n}} \tag{22}$$

For varying sample sizes n_j, the centre line (CL) is computed as above.
The control limits for sample j are given as

$$\text{UCL} = \bar{p} + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{p}(1-\bar{p})}{n_j}} \tag{23}$$

$$\text{LCL} = \bar{p} - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{p}(1-\bar{p})}{n_j}} \tag{24}$$

where $p_j = d_j / n_j$ is the proportion of defectives in batch j.

The C-Chart

The C-chart assumes that defects in a quality attribute are rare, therefore, the control limits are computed based on the Poisson distribution. The chart is constructed by plotting the number of defectives per batch. The statistics involve in the construction of the C-chart are as follow:

$$\text{CL} = \bar{c} = \frac{1}{m}\sum_{j=1}^{m} c_j \tag{25}$$

where c_j is the number of defects in unit/batch j, and m is the number of units inspected. The control limits are

Tibiscus

$$\text{UCL} = \bar{c} + Z_{1-\frac{\alpha}{2}}\sqrt{\bar{c}} \tag{26}$$

$$\text{LCL} = \bar{c} - Z_{1-\frac{\alpha}{2}}\sqrt{\bar{c}} \tag{27}$$

The U-Chart

Unlike the C chart, the U-chart does not require a constant number of units, and it can be used, for example, when the batches are of different sizes. In this chart we plot the *rate of defects*, that is, the number of defects divided by the number of units inspected.

The centre line for a U-chart is computed by

$$\text{CL} = \bar{u} = \frac{1}{\sum_{j=1}^{m} n_j}\sum_{j=1}^{m} d_j \tag{28}$$

where d_j is the number of defects in batch j, and n_j is the sample size in batch j. The control limits are

$$\text{UCL} = \bar{u} + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{u}}{n}} \tag{29}$$

$$\text{LCL} = \bar{u} - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{u}}{n}} \tag{30}$$

For varying sample sizes, the centre line is computed as above. However, the control limits for sample/batch j are given as

$$\text{UCL} = \bar{u} + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{d_j}{n_j}} \tag{31}$$

$$\text{LCL} = \bar{u} - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{d_j}{n_j}} \tag{32}$$

where d_j is the number of defects in sample/batch j. Note that $\frac{1}{n_j}d_j$ will be plotted.

3.3 *The Process Capability Indices*

Process capability compares the output of an *in-control* process to the specification limits by using *capability indices*. A process capability index ([ZSW90, WB91]) uses both the process variability and the process specifications to determine whether the process is "capable". A number of statistics are used to measure the capability of a process. This include C_p, C_{pk} and C_{pm} as defined below:

$$C_p = \frac{USL - LSL}{e\sigma} \tag{33}$$

$$C_{pk} = \min\left(\frac{USL - \overline{X}}{e\sigma}, \frac{\overline{X} - LSL}{e\sigma}\right) \tag{34}$$

$$C_{pm} = \frac{USL - LSL}{e\sqrt{\sigma^2 + \left(\overline{X} - t\right)^2}} \tag{35}$$

where USL is the upper specification limit, LSL is the lower specification limit, t is the target value and $\overline{X}$ is the mean of the process performance obtained from the sample. A value of 6 is often used for e in expressions (33) and (35) while 3 is used for it in expression (34) ([Bis90]). Also, the estimate of σ, the standard deviation of the process, is obtained using the group ranges as $\frac{1}{m}\sum_{j=1}^{m}\frac{R_j}{d_2(n_j)}$ or the standard deviation as $\frac{1}{m}\sum_{j=1}^{m}\frac{S_j}{c_4(n_j)}$ or the pooled variances as $\frac{S_p}{c_{4\left[\sum_{j=1}^{m} n_j - m + 1\right]}}$ where both $d_2(n_j)$ and $c_4(n_j)$ are as earlier defined for equation (5) and $S_p = \frac{1}{k-m}\sum_{j=1}^{m}(n_j - 1)s_j$ with $k = \sum_{j=1}^{m} n_j$.

4. APPLICATION OF THE SQC SOFTWARE

This section presents how to use SQC package to analyse various forms of quality control data and construct the necessary control charts.

A double click on ”Shortcut to Super QC-pack” located on the desktop or on program menu will prompt the user interface window given in Fig 1. This is the parent window where the user can select the type of task he wants to perform. Here, the user can click on “CLICK HERE TO SELECT TYPE OF CHART” panel to select from any of the seven different types of control charts if he intends to construct control chart or select the “Process Capability” if he intends to estimates the various forms of process capability indices.

Tibiscus

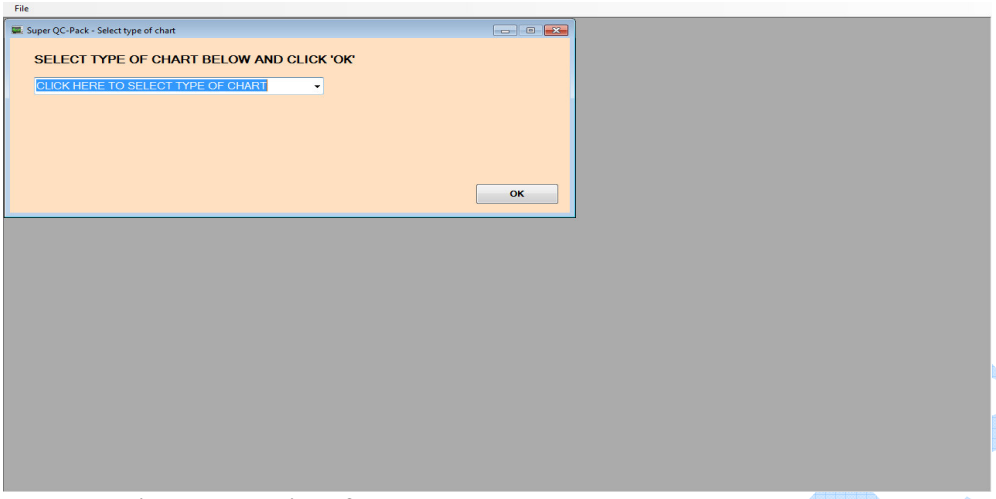

Figure 1: The parent window or the interface where the user can select the type tasks he wants to perform. The available options are the C, P, R, S, X-bar and U charts as well as 'Process Capability' to construct all the aforementioned control charts as well as estima0ting some of the process capability indices

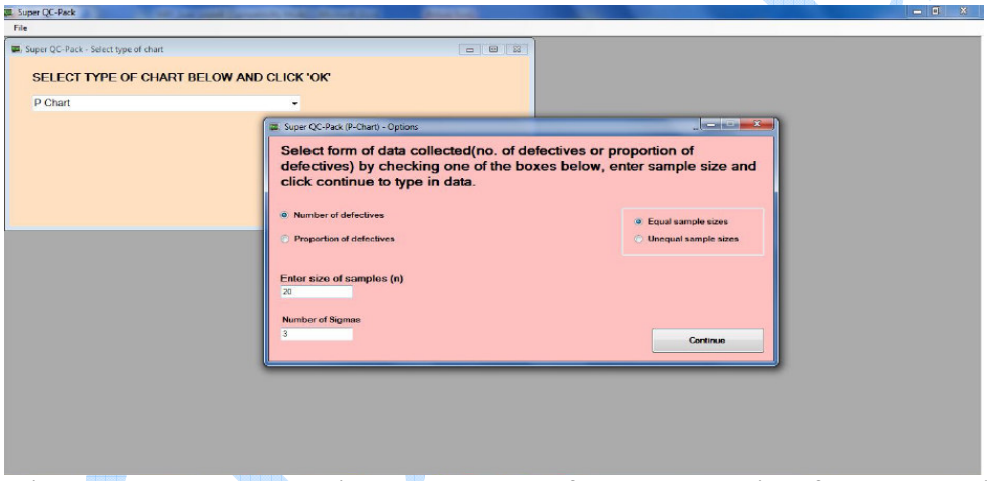

Figure 2: The window where some task options can be selected for the construction of P-chart. In this illustration, the sample size is 20, the number of sigma is 3, the samples have equal sizes, and the form of data to be supplied is the number of defectives per sample (raw data)

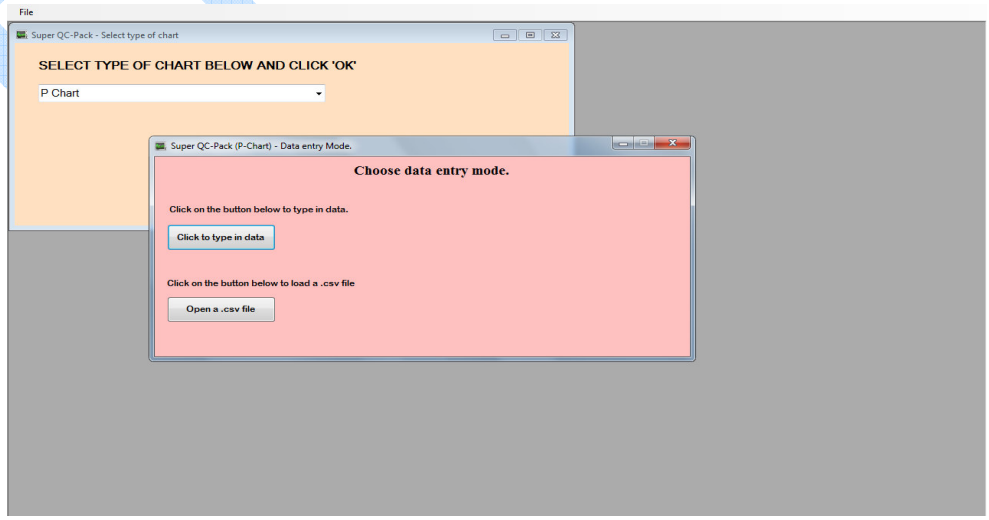

Figure 3: An interface where method of providing data for analysis. This window interface shows the option for imputing data directly by checking the box "Check to type in data" and the option for uploading a comma separated values (csv) data file by checking the box "Open a .csv file"

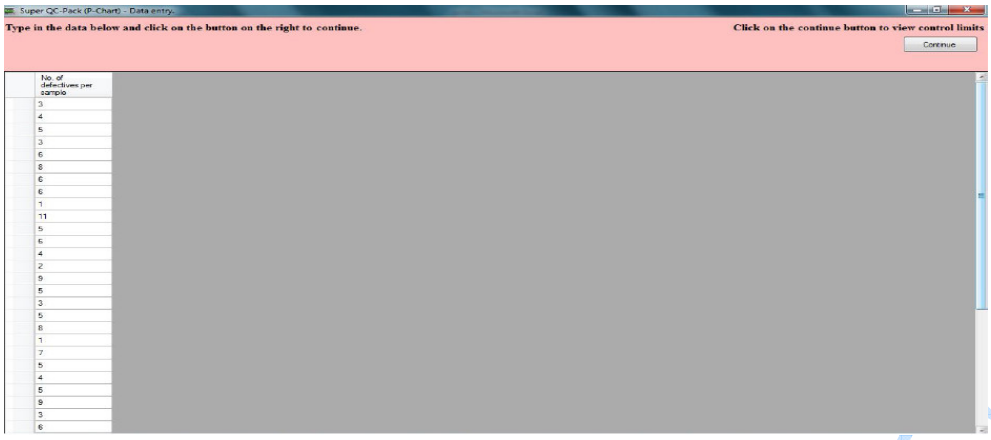

Figure 4: The data entry window interface in which the numbers of defectives per group have been entered for forty groups/batches with 20 samples per batch

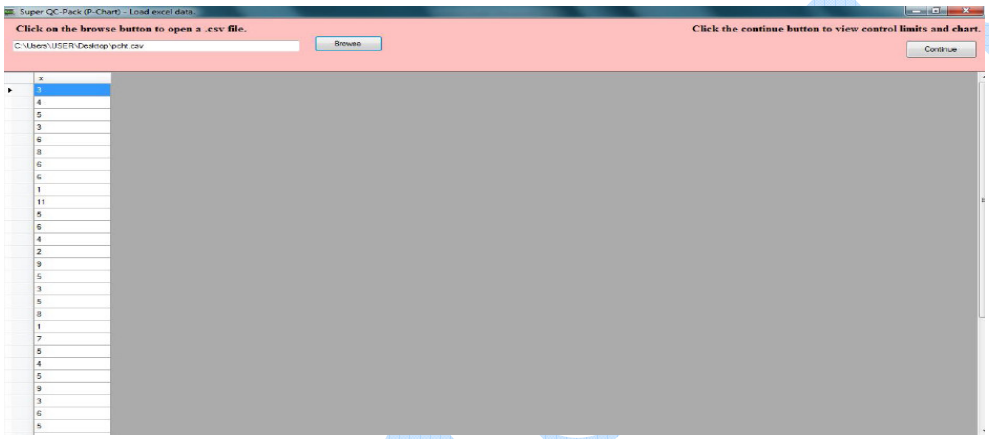

Figure 5: The window interface through which an excel CSV data file containing the number of defectives for forty batches with 20 samples per batch was uploaded into the SQC environment for analysis

As an illustration, if the P-chart is selected and "OK" icon is clicked for instance, another interface will be displayed where further options can be selected depending on the form of the data to be analyzed as in Fig 2. Here, the user have to select "Number of defectives" or "Proportion of defectives" option if the data to use are the number of defectives (raw data) or the proportion of defectives (summary statistics from data) in each batch respectively.

If the sample sizes are equal across all the groups, the "Equal sample size" option should be selected and such sample size should be entered in the "Enter size of sample (n)" box. However, if the sample sizes are not equal across the groups, the 'Unequal sample sizes' option should be checked and the respective sample size in each group should be part of the data to be typed in or uploaded for analysis. Finally, the number of sigma that determines the upper and lower specification limits should put into "Number of sigma" box. The default is 3 and this can be varied depending on user's interest. Click on "Continue" to proceed to the next window where the method of providing data shall be chosen as shown in Fig 3.

In the data entry mode interface (Fig 3), the user should choose whether he wants to type in the data at the next data entry interface by checking the "Click to type in data" box or upload a comma separated values (CSV) data file by checking the "Open a .csv file" box.

If "Click to type in data" box is selected, one column data interface will be displayed in which to enter the number of defectives or proportion defectives (depending on the form of the data) for equal sample sizes as demonstrated in Fig 4. For unequal sample sizes, additional column will appear where the sample size for each group will be entered.

If the "Open a .csv file" box is clicked, another interface will appear that allows the user to browse and locate the directory where the excel CSV data file to upload is resident as shown in Fig 5. For equal sample size per group, the excel CSV file should have one column containing the number of defectives or the proportion defectives (depending on the available data) with the first row being the column label. For unequal sample sizes, additional column should be included to contain the sample sizes for respective groups

After the data have been either entered manually or uploaded as shown in Fig 4 or Fig 5, a click on "Continue" icon will produce the p-chart as shown in Fig 6. The chart in Fig 6 shows the plot of proportion of defectives per group/batch for all the twenty batches. The estimated upper and lower control limits and the center line were drawn and reported on the chart. Another interesting feature of the constructed p-chart by SQC software is that the estimates of the center line, upper and lower control limits of the corresponding np-chart are simultaneously reported on any p-chart constructed as shown in Fig 6.

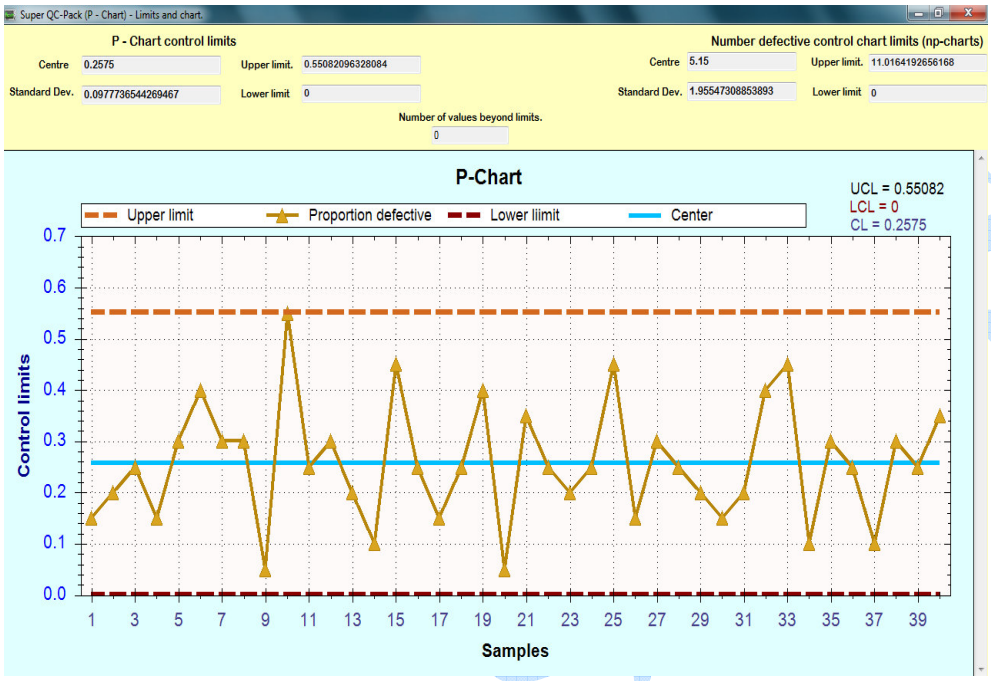

Figure 6: The graphical output of SQC software which is typical P-chart constructed using data set on forty batches at equal sample size of 20. The proportion of defectives per sample was plotted for the various samples. The graph reports the center line estimate, the upper and lower control limits and a space that reports the number of point(s) beyond these limits (if any). The estimates of the center line, upper and lower control limits for the corresponding np-chart are equally reported on the p-chart

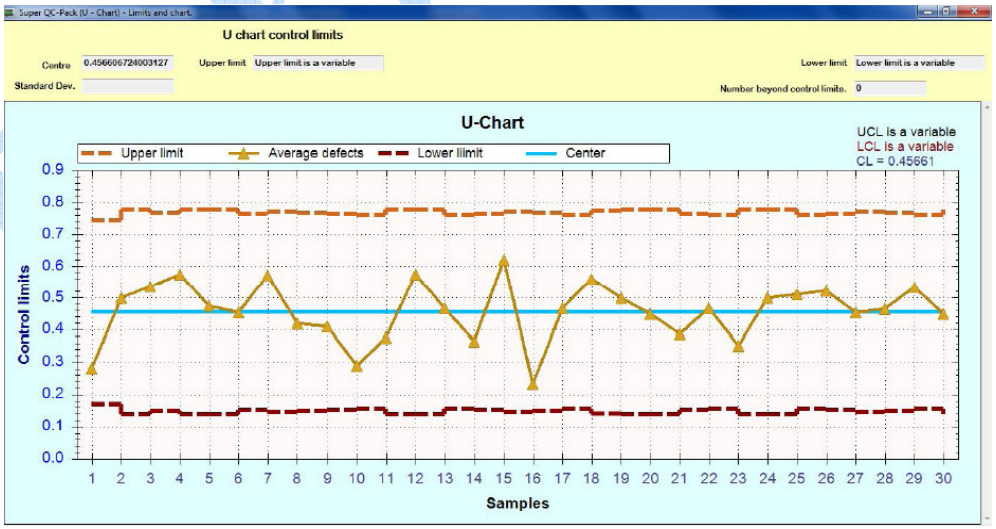

Figure 7: A typical U-chart output of SQC software for 30 batches/samples with unequal sample sizes across batches

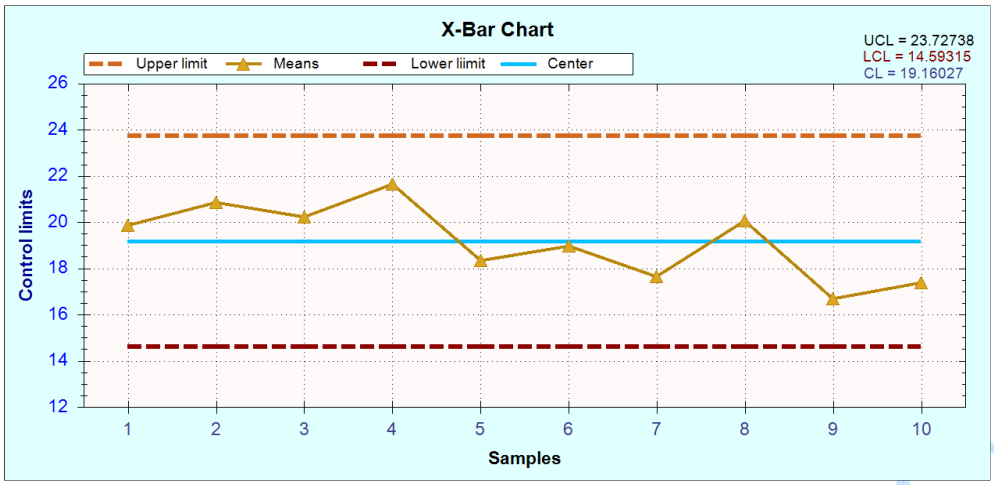

Figure 8: A typical X-bar chart output of SQC software for 10 batches/groups with equal sample sizes across batches

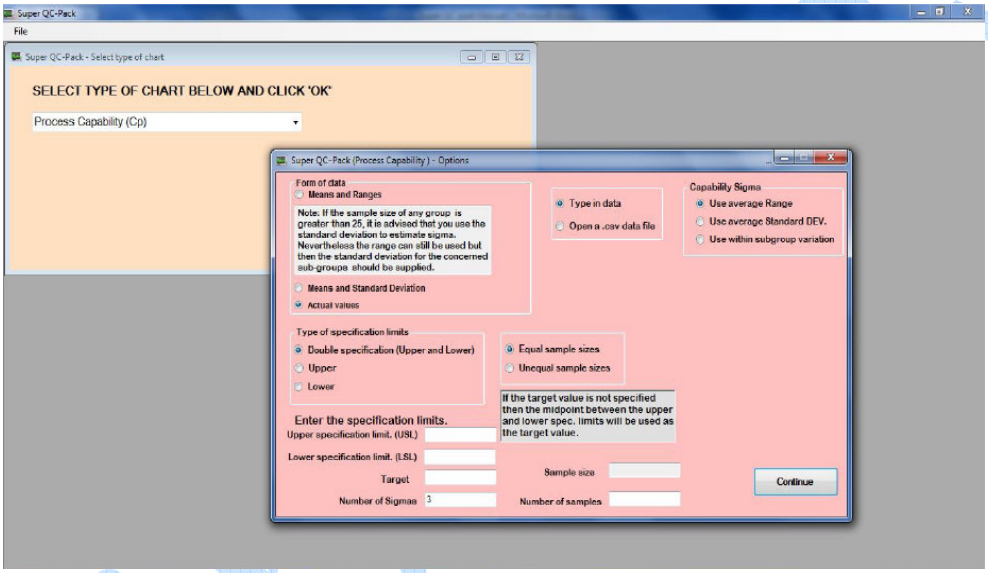

Figure 9: The SQC window interface where all the necessary statistics can be supplied to estimate the required process capability indices

More generally, other control charts can be easily constructed with the SQC software by checking the necessary boxes till the final graph is displayed. More details on this are provided in the user's manual that accompanied the software. Nonetheless, we provide typical U-chart and X-bar chart outputs of the SQC software in Fig 7 and Fig 8 to demonstrate their graphical outlook.

The U-chart displayed in Fig 7 was developed by plotting the proportion of defects in each batch/sample against the batch/sample number with unequal sample sizes across all the batches. Except the center line, the estimates of upper and lower control limits were declared to be variable on the chart since these statistics varied and were computed for each batch.

The X-bar chart in Fig 8 displays the center line, upper and lower control limits for ten samples/batches with equal sample size.

Finally, various process capability indices such as C_p, C_{pk} and C_{pm} can be computed for a manufacturing or service process by supplying the necessary statistics into the SQC environment like the process mean, the upper and lower specification limits and the target value. To do this, simply launch the SQC package through the shortcut, select "Process Capability (Cp)" and click 'OK' to continue. A window interface where all the required statistics can be supplied and in which all necessary options can be selected, as shown in Fig 9, will be displayed.

Some of the options displayed in Fig 9 allow the user to select if the raw data ("Actual values") or their summary statistics ("Means and Ranges" or "Means and Standard Deviations") are to be used.

If "Actual values" options is selected, the method of estimating capability sigma among "use average Range", "Use average Standard Dev." or "Use within subgroup variation" option has to be selected depending on user's preference. With this option, the

number of samples/batches has to be provided if the data set has to be supplied manually through "Type in data" option. Indicate whether the group sample sizes are equal or not by selection the appropriate option. Lastly, the user should provide the upper specification limit (USL), lower specification limit (LSL) and the mean of the process performance. Also, the target value has to be provided to compute the value of C_{pm}. Indicate the number of sigma to be used in the relevant box and click "Continue" to type in or upload data set for analysis.
If any of the summary statistics ("Means and Ranges" or "Means and Standard Deviations") is selected, the user has to select the group sample size for equal sample size situation. Further details can be found in the SQC manual.

5. CONCLUDING REMARKS

A novel statistical package, the Super QC-Pack version 1.0.0, for analyzing quality control data has been introduced in this paper. This software is user-friendly that can be used by quality control officers with limited knowledge of statistics and/or statistical computing.
Some of the existing statistical packages were designed to accept only raw data sets inform of the actual measurements of quantities that emanated from the manufacturing or service processes. However, in many situations, only the summary statistics of the raw data sets like the mean, the range, the standard deviations and so on might be the only available information with which the stability of the processes should be assessed. Therefore, one of the benefits of the new SQC package over some of the existing statistical packages like SPSS ([***11]), MINITAB, and the like for analyzing quality control data is that it permits the use of both raw data and summary statistics of data generated by any manufacturing or service processes.
Another advantage of the SQC package over others lies in its ability to compute process capability indices (PCI) for products or processes that are characterized with single specification limit (i.e. the smaller-the-better and the larger-the-better types of characteristics). Some of the existing statistical packages like SPSS, MINTAB computes PCI only for double specification limits in which both the USL and LSL have to be provided.
The SQC package has been thoroughly evaluated and tested over different platforms, and has been found to produce accurate computational results with excellent graphical outputs. This software has been equally validated and found to meet Installation Qualification (IQ), Operational Qualification (OQ) and Performance Qualification (PQ) (IQ/OQ/PQ) criteria, Revolutions ([***09]).
The SQC version 1.0.0 is accompanied with a detailed manual that provides a step-by-step of how to use the software to analyse quality control data. Not only this, the manual also chronicles how the required estimators for the construction of various control charts were developed. However, it is important to remark that this manual is not to be substituted for standard statistical textbooks that specifically address the theoretical formulations of various procedures for constructing the various charts discussed here.
Without loss of generality, studies have shown that all the existing statistical packages produce computational results with varying degrees of accuracies. However, the development of any new statistical packages to perform specific tasks has always been an efficient tool to validate and improve on the functionalities and results of the existing packages. This position is further justified through a quick comparison of the results of the SQC package to the results produced by SPSS 16, R (version 2.14.1) and MINITAB 15 statistical packages for the construction of X-bar and R-charts as given in Table 1 and Table 2 using a simulated data set of a typical manufacturing process. The analysis in R package was performed through the qcc library ([Scr04]).

Table 1: The results of the new SQC, SPSS, R and MINITAB statistical packages showing the estimates of the center line(CL), upper control limit (UCL) and lower control limit (LCL) for the construction of the X-bar chart. The exact results provided by analytical method are equally reported. All results are reported up to 4 decimal places.

Packages	LIMITS		
	CENTER	UCL	LCL
SPSS	9.8441	12.7390	6.9492
R	9.8441	12.7389	6.9492
MINITAB	9.8440	12.7390	6.9490
SQC-PACK	9.8441	12.7386	6.9496
Exact results	9.8441	12.7386	6.9496

Ten production batches were conjectured in our data simulation and twenty samples of a given unit were simulated per batch. These simulated data are presented in the Appendix. To have equal platform for results comparisons, all the results were reported up to four decimal places as shown in Tables 1 and 2. From the results these tables, it can be observed that the new SQC statistical package provides estimates of the CL, UCL and LCL that agreed perfectly with the exact results in both the X-bar and R-charts. Results of SPSS, R and MINITAB packages equally showed appreciable agreements with the exact results, though with minor deviations from the exact results as shown on the two tables.

These results further justify the need for constant appraisal of some of the existing statistical packages for their possible improvements to meet emerging challenges.

Table 2: The results of the new SQC, SPSS, R and MINITAB statistical packages showing the estimates of the center line(CL), upper control limit (UCL) and lower control limit (LCL) for the construction of the R-chart. The exact results provided by analytical method are equally reported. All results are reported up to 4 decimal places.

Packages	LIMITS		
	CENTER	UCL	LCL
SPSS	16.1180	25.5519	6.6842
R	16.1181	25.5519	6.6842
MINITAB	16.1200	25.5500	6.6800
SQC-PACK	16.1180	25.5507	6.6854
Exact results	16.1180	25.5507	6.6854

Unlike some of the existing commercialized statistical packages, the new SQC package is free software which can be downloaded at www.unilorin.edu.ng/sqcpack/yahya_adeniyi/webpage.htm as earlier remarked.
Apart from some other useful features the Super QC-pack software has regarding its functionality, the package has a very rich graphical interface that easily produces high quality graphical outputs of the intended control charts. Not only this, the SQC package is computationally efficient and provides accurate results like any of the existing statistical packages.

6. FUTURE DEVELOPMENT

The first version of the new SQC software (SQC 1.0.0) is presented in this paper. However, additional features shall be incorporated in the subsequent versions of the software to improve its current performance. Nonetheless, few additional features that will reflect in the second version of the software (SQC 1.0.1) are the implementation and construction of the variance (s^2) chart, cumulative sum (CUSUM), and exponentially weighted moving average (EWMA) charts ([WA93]) among others.
Users of the current version of the SQC software (SQC 1.0.0) are therefore enjoyed to forward their useful comments and observations to the developers to improve the subsequent versions of the package.

REFERENCES

[Bis90] **A. F. Bissell -** How Reliable is your Capability Index?. *Applied Statistics*, 39, 331-340, 1990.

[B+07] **C. Bauman, J. De Heck, E. Leonard, M. Miranda -** SPC: Basic control charts: theory and construction, sample size, x-bar, r charts, s charts., https://controls.engin.umich.edu/wiki/index.php/SPC:_Basic_control_charts:_theory_and_construction,_sample_size,_x-bar,_r_charts,_s_charts, 2007, Accessed on 27th April 2012.

[CW87] **C. W. Champ, W. H. Woodall -** Exact results for Shewhart Control Charts with Supplementary runs rules, *Technometrics*, 29, 393-399, 1987.

[Dun86] **A. J. Duncan**. *Quality Control and Industrial Statistics*, 5th ed., Irwin, Homewood, IL, (1986).

[DRT12] **H. Deng, G. Runger, E. Tuv -** System monitoring with real-time contrasts. *Journal of Quality Technology,* 44(1), 9-27, 2012.

[ET02] **B. Eric, R. Tomy**. *Your visual blueprint for developing .NET applications*. Hungry Minds, Inc., 2002.

[KJ92] **S. Kotz, N. L. Johnson**. *Process Capability Indices*, Chapman & Hall, London, 1992.

[LS90] **J. M. Lucas, M. S. Saccucci -** Exponentially weighted moving average control schemes: Properties and enhancements, *Technometrics 32*, 1-29, 1990.

[Mon04] **D. C. Montgomery -** *Introduction to Statistical Quality Control*. 4th edition. John Wiley & Sons Inc., New York, NY, USA, (2004).

[Mon05] **D. C. Montgomery -** *Introduction to Statistical Quality Control*. 5th edition. John Wiley & Sons Inc., New York, NY, USA, 2005.

[Mon85] **D. C. Montgomery -** *Statistical quality control*. New York: Wiley, 1985.

[Mon91] **D. C. Montgomery -** *Design and analysis of experiments*, 3rd ed. New York: Wiley, 1991.

[OS90] **E. R. Ott, E. G. Schilling -** *Process Quality Control*, 2nd ed., McGraw-Hill, New York, NY, 1990.

[Que93] **C. P. Quesenberry -** The effect of sample size on estimated limits for $\bar{X}$ and **X** control charts, *Journal of Quality Technology*, 25(4) 237-247, 1993.

[Rya00] **T. P. Ryan -** *Statistical Methods for Quality Improvement*, 2nd ed., Wiley, New York, N, 2000.

[RS97] **T. P. Ryan, N. C. Schwertman -** Optimal limits for attributes control charts. *Journal of Quality Technology*, 29 (1), 86-98, 1997.

[Scr04] **L. Scrucca -** qcc: an R package for quality control charting and statistical process control. *R News* 4/1, 11-17, (2004).

[Wil06] **L. Wilkinson** . Revising the Pareto Chart. *The American Statistician*, 60: 332–334, 2006.

[Woo97] **W. H. Woodall -** Control Charting Based on Attribute Data: Bibliography and Review, *Journal of Quality Technology*, 29, 172-183, 1997.

[WA93] **W. H. Woodall, B. M. Adams -** The Statistical Design of CUSUM Charts, *Quality Engineering*, 5(4), 559-570, 1993.

[WB91] **G. B. Wetherill, D. W. Brown -** *Statistical Process Control*. New York: Chapman & Hall, 1991.

[ZSW90] **N. F. Zhang, G. A. Stenback, D. M. Wardrop -** Interval Estimation of the Process Capability Index, *Communications in Statistics: Theory and Methods,* 19(21), 4455-4470, (1990).

[***09] **Revolutions**. Analyzing clinical trial data for FDA submissions with R. http://blog.revolutionanalytics.com/2009/01/analyzing-clinical-trial-data-with-r.html , 2009, (accessed on April 30, 2012).

[***11] SPSS Inc., SPSS Base 18.0 for Windows User's Guide. SPSS Inc., Chicago, 2011. www.spss.com

[***a] Minitab Inc. Minitab 15, State College PA 16801-3008, (USA) http://www.minitab.com

[***b] R Development Core Team. R: A language and environment for statistical computing. R foundation for Statistical computing, Vienna, Austria, (2012), ISBN 3-900051-07-0, URL *http://www.r-project.org/*.

Tibiscus

APPENDIX

Simulated data of a typical manufacturing process containing production outputs from m = 10 batches with equal sample size n = 20 per batch. This data set is suitable for variable control charts.

Batch 1	Batch 2	Batch 3	Batch 4	Batch 5	Batch 6	Batch 7	Batch 8	Batch 9	Batch 10
6.858599	3.592168	6.581251	14.28892	9.417935	8.048522	11.66762	12.78362	13.96662	0.800209
4.567071	13.82091	15.63457	2.463062	13.75399	2.241677	10.14516	7.925387	8.743445	13.11155
10.75741	5.659864	10.17501	11.27845	6.52225	10.57861	26.07716	12.56768	10.89477	10.77035
3.822621	4.326155	14.05839	11.26844	10.14711	5.532482	2.621642	7.429197	16.42043	9.137086
8.252853	6.660277	15.04603	9.78798	4.075502	3.409243	14.16487	8.240209	10.79412	18.83831
13.06703	9.789829	9.414519	14.08896	12.68272	9.904515	3.470926	7.917569	15.63792	9.589729
8.211388	7.052706	10.15898	6.156379	13.48273	9.670436	14.63169	5.953594	5.869654	0.519684
11.64456	14.58843	8.970714	12.1662	4.898261	5.640258	4.83789	15.02356	2.178464	5.580832
10.9385	13.29911	7.184101	7.579476	13.21604	13.57501	13.80031	10.39023	8.182295	13.5236
6.997874	12.4356	15.54431	3.318781	3.021712	8.852816	16.26958	16.06914	8.359524	8.498562
18.09912	11.75001	13.24593	14.7615	11.64341	9.540027	15.78168	7.897141	5.75663	-0.81868
13.47377	3.858101	12.12054	14.8196	8.424936	9.781298	11.41188	7.12973	9.845145	0.786026
8.990083	9.759029	7.473844	11.37529	6.632392	11.80618	9.318151	12.45215	8.432307	11.97941
2.82170	6.412119	8.631535	14.53187	15.64609	11.86221	7.787571	12.26256	-0.0338	20.26574
7.70490	8.316696	12.45317	10.68339	8.93358	16.71965	9.389673	7.871844	11.67467	12.74362
9.043803	4.453024	2.100986	7.921679	11.45497	3.412445	14.89345	20.98837	9.225605	17.22916
10.94766	10.92494	15.16175	4.503849	13.36279	15.80347	12.7332	12.70971	12.06283	5.414147
1.992735	7.327522	9.854177	5.050078	3.493924	6.226976	20.66991	3.905444	9.755842	16.0271
7.957671	9.925048	4.58413	7.450863	16.51353	10.05773	20.96741	16.40317	6.814633	9.487041
11.78846	9.42226	13.63299	0.322238	10.97089	12.45052	9.362633	10.57516	13.83587	-0.14718