العنوان:	Using Multivariate Statistical Quality Control Models to Monitor the Quality of Drinking Water in Khan Younis Governorate - Palestine
المصدر:	المجلة العربية للجودة والتميز
الناشر:	مركز الوراق للدراسات والأبحاث
المؤلف الرئيسي:	Al-Telbany, Shady Ismail
مؤلفين آخرين:	Ashour, Mohammed Fayez(Co-Auth)
المجلد/العدد:	مج2, ع2
محكمة:	نعم
التاريخ الميلادي:	2015
الشهر:	مايو
الصفحات:	5 - 32
رقم MD:	793173
نوع المحتوى:	بحوث ومقالات
اللغة:	English
قواعد المعلومات:	EcoLink
مواضيع:	المياه، مياه الشرب، الجودة الإحصائية، النماذج الإحصائية، محافظة خان يونس، فلسطين
رابط:	http://search.mandumah.com/Record/793173

Using Multivariate Statistical Quality Control Models to Monitor the Quality of Drinking Water in Khan Younis Governorate -Palestine

Dr. Shady Ismail Al-Telbany
AL-azhar University-Gaza
Mohammed Fayez Ashour

Abstract

Groundwater is one of the most precious natural resources in the Gaza

Strip as it is the only source of drinking water for the majority of the population. So aim of this study is to evaluate the statistical methods that are used to monitor the quality of drinking water in order to suggest the best statistical models that are used in monitoring and detecting small changes to avoid diseases that may be caused by the problem of water pollution. Our data set were taken for several readings of groundwater wells from Khan Younis governorate for three variables which is the major chemical components of drinking water, which is mainly in judging the quality of drinking water, namely: (Chloride (CL), nitrate (NO3), total dissolved salts in water (TDS)), during period from 1987 to 2012.

In this study, univariate control models (Shewhart, EWMA, CUSUM) were applied to the same data set , then make a comparison between three models but was reached that univariate control models have not achieved good control in the detection of small changes, as well as when we resort to explain the problem of variable we need to read and interpret more than one model ,so it has been applied multivariate control models (Hotelling, MEWMA, MCUSUM) on the same data set and found that it is more sensitive in detecting small changes from than univariate control models, because it give one explanation and one read for more than one variable, then make a comparison between the three types to find the best control model to monitor the quality of drinking water and detect small changes. It was concluded that the MCUSUM model is the best and fastest in achieving better quality control and detect small changes to monitor the quality of drinking water.

1. Introduction:

Quality control (QC) is an important function in factory as it deals with products inspection before the product was shipped to customers. Therefore, quality has become one of the most important consumer's decision factors in the selection among competing products and services. The using of statistical quality control (SQC) techniques to manufacturing become an important topics of study for a lot of research.

Statistical process control (SPC) is a powerful collection of problem solving tools useful in achieving process stability and improving capability through the reduction of variability. SPC is one of the greatest technological developments of the twentieth century because it is based on sound underlying principles, is easy to use, has significant impact, and can be applied to any process.

The main work of Statistical Quality Control is to control the central tendency and variability of some processes, a common monitoring tool is to construct control models (Dou and Ping, 2002). A control model is a statistical scheme (usually allowing graphical implementation) devised for the purpose of checking and then monitoring the statistical stability of process. The most widely used method to control the central tendency of a process is Shewhart-X model "(Shewhart 1931)" which includes a centerline and two control limit lines.

There are two other possible alternatives to the Shewhart control models in the construction of the central location control models. One is the CUSUM (Cumulative Sum) model and the other is the EWMA (Exponentially Weighted Moving Average) model. Both of these concentrate on improving the performance of control models in detecting small shifts by using historical data (Dou and Ping 2002).

Univariate statistical process control models, mainly Shewhart model, Cumulative sum (CUSUM) & EWMA control models, have received considerable attention in industry due to their case of use by the production personnel and others with minimal statistical knowledge. However, a USPC models can only one variable at a time, which means that process engineers have to look at fifty control models to monitor the process. Furthermore, those univariate models do not take any possible correlation among variables into account. Monitoring of process in which several related variables are of interest is collectively known as multivariate statistical process control

(MSPC). MSPC is a methodology, based on quality control models, that is used to monitor the stability of a multivariate process.

In modern manufacturing environments, the characteristics or variables of a multivariate process often are interrelated and from a correlated set. Since the variables do not behave independently, they must treat together as a group, and not separately. Particularly in the chemical industry, where input is being chemically altered to produce a particular output, the variables of interest are usually the components produced by the previous process.

Stability is achieved when the means, variance, and covariance of the process variables remain stable over rational subgroups of the observations. The conventional MSPC chars mainly include multivariate Shewhart control models, multivariate CUSUM control models, multivariate EWMA control models.(Zhao, 2007).

Jackson stated that any multivariate process control procedure should fulfill four conditions: a) an answer to the question: "Is the process in control?" must be available; b) an overall probability for the event "Procedure diagnoses an out-of-control state erroneously" must be specified; c) the relationships among the variables - attributes should be taken into account; d) an answer to the question: "If the process is out-of-control, what is the problem?" should be available. The Jackson's fourth condition is the most challenging problem at this time in the MSPC area, an appealing subject for many researchers in the last years, and the main topic under consideration in this article.

As it obvious, there is a lot of concern by many researchers and producers to study the quality of drinking water because water pollution is a major global problem which requires ongoing evaluation and revision of water resource policy at all levels (international down to individual aquifers and wells). It has been suggested that it is the leading worldwide cause of deaths and disease, for this reason the quality of drinking water must be controlled.

The Multivariate Statistical Quality control Models such as Hotelling's model, Multivariate cumulative sum (MCUSUM) model and Multivariate Exponential Weighted Moving Average (MEWMA) model is an important statistical process tools for analysis and monitoring the quality of drinking water.

2. Study Problem

In the last period water pollution is a major global problem which requires ongoing evaluation, which resulted from sea water and sewage leakage on groundwater wells. For this reason, the problem of the present study is how to evaluate the statistical methods and propose better statistical tools to monitor the quality of drinking water in khan younis governorate.

3. Objectives of the Study

1). Demonstrate and compare the effectiveness of three univariate statistical models (EWMA &CUSUM) to determine which is more effective in detecting small shifts.

2). Demonstrate and compare the effectiveness of three multivariate statistical quality control models (Hotelling T^2 , MEWMA , MCUSUM) to determine which model is the useful, sensitive and adequate in identifying changes or shifts in the quality of water.

3). Monitor and evaluate the state of the quality of drinking water to detect if there is a deviation from target line specification (standard specification) or if there is an excessive variability around it by using three important statistical control models (Hotelling T^2, MEWMA, MCUSUM).

4. Study Methodology

The models which used in this study cover the use of Hotelling, MEWMA & MCUSUM techniques with commercially available software to monitor the quality of drinking water in Khan Younis governorate. However, the finding may be equally as applicable to any type of process under statistical control for which the reader desires to know if the drinking water data is changing. For the overall proposed methodology, the author presents an approach to monitor the quality of drinking water using (Hotelling, MEWMA, MCUSUM) control models and compared the results.

Since the quality control models have long tradition in engineering and biological research most of the terminology coincides with the one that is used in these fields such as: concrete blocks, performance of hospitals quality of water and etc. Quality control techniques have been applied to monitor the quality of drinking water.

The CUSUM & EWMA will be explored to determine the best conditions for monitoring quality of drinking water, the performance of the CUSUM model and EWMA model is compared to determine which model is the useful, sensitive and adequate in identifying changes or shifts in the process. As well, the Hotelling model, MEWMA model & MCUSUM model will be explored and compared to determine which model is more sensitive and useful if monitoring the quality of drinking water. Furthermore, we will compare between (EWMA & MEWMA, CUSUM & MCUSUM) to decide which is more sensitive in detecting small shifts.

5. Multivariate Statistical Quality Control Models

Statistical process control is based on a number of basic principles which apply to all processes, including batch and continuous processes of the type commonly found in the manufacture of bulk chemicals, pharmaceutical products, specialist chemicals, processed foods and metals. The principles apply also to all processes in service and public sectors and commercial activities, including forecasting, claim processing and many financial transactions. One of these principles is that within any process variability is inevitable (Chanda, 2001).

Generally there are two groups of (SPC), i.e. univariate statistical process control (USPC) and multivariate statistical process control (MSPC), which are used for different scenarios. The process of monitoring and control primarily apply to the systems or processes from the univariate perspective, which has only one process output variable or quality characteristic measured and tested. If a process is to meet or exceed customer expectations, generally it should be produced by a process that is stable or repeatable. More precisely, the process must be capable of operating with little variability around the target or nominal dimensions of the producer's quality characteristics.

Typically process monitoring applies to systems or processes in which only one variable is measured and tested. There are many processes in which the simultaneous monitoring or control of two or more quality characteristics is necessary. Process monitoring problems in which several variables are of interest are called (MSPC). One of the disadvantages of a univariate monitoring scheme is that for a single process, many variables may be monitored and even controlled. MSPC methods overcome this disadvantage by monitoring several variables simultaneously. Using multivariate statistical process control methods, engineers and manufacturers who monitor complex processes may monitor the stability of their process.

The first original study in multivariate quality control was introduced by Hotelling (1947) (Runger & Montgomery, 1997).

An important aspect of the Hotelling's-T^2 Control model is how to determine the sample variance-covariance matrix used in the calculation of the model statistics (UCL and LCL). When rational subgroups are taken, the implication is that the appearance of a special cause of variation within a subgroup is unlikely, so that all observations within a subgroup share a common distribution. Thus, the regular sample variance-covariance matrix is useful and taking the average over all the subgroups is the common procedure, unless there are special causes that alter the variance-covariance matrix. If subgroups are taken and the population parameters are known then the Hotelling's T^2 statistic, T_i^2, is $\chi^2_{\alpha,p}$ distributed, where p is the number of variables and α is the probability of false alarm. In the event that the population parameters are unknown (that is, the mean vector and the variance-covariance matrices are unknown), the estimates are obtained from the sample and the Hotelling's T^2 statistic, T_i^2 has an F or Beta distribution (Kolarik, 1999).

5.1 The Multivariate Normal Distribution

The multivariate normal distribution is the core of the multivariate statistical analysis. This is due to the fact that the sampling distribution of multivariate distributions exhibit approximately normality due to the central limit theorem in the univariate case if a random variable is normally distributed with mean μ and variance σ^2 it has a density function (Santos, 2012).

5.2 Types of Multivariate Statistical Quality Control

Since its introduction by W. A. Shewhart, a physicist and statistician working for Bell laboratory, the Shewhart control model has become a popular tool for monitoring the performance of industrial processes. Montgomery (1996) gave a detailed research of Shewhart control model.

The same is true with the (CUSUM) control model proposed by Page (1954) and Exponentially Weighted Moving Average (EWMA) control scheme proposed by Roberts (1959).

Hawkins and Olwell (1998) gave a comprehensive and systemic description of CUSUM model while Lucas and Saccucci (1990) presented a detailed research of the properties of EWMA model.

SHEWHART, CUSUM and EWMA schemes are acknowledged as the most widely used control models.

In modern quality control, it is becoming common to monitor several quality characteristics of a process simultaneously. This challenge motivates attempts to extend the univariate Shewhart, CUSUM and EWMA statistics to multivariate data. In the past decades, several kinds of multivariate control model for the process mean have appeared, most of them are generalizations of their corresponding univariate procedures.

Three of the most useful multivariate quality control statistics are Hotelling's T^2 (Hotelling, 1947), MCUSUM proposed by Woodall and Ncube (1985), MCUSUM proposed by Crosier (1988) and MEWMA proposed by Lowry (1992).

The classical application of these three types of control schemes, namely, that the process being inspected follows a multivariate normal distribution. (Dai, et al., 2009)

- **Hotelling T^2 control Models**

In many industrial settings it is frequently required to monitor more than one interrelated variables. The Hotelling's T^2 control model is one of the multivariate statistical tools which are widely used to detect the presence of special-causes of variation by monitoring a mean vector μ. This model is popular as it possesses almost all the desirable characteristics for a multivariate control model such as ease of application, flexibility, sensitivity to small process changes, and the availability of software for application (Mason & Young, 2002).

Like any other control models for monitoring the variability in a process, its construction consists of Phase I and Phase II (Alt, 1985) which are also referred to as retrospective and prospective analysis respectively (Woodall & Montgomery, 1999). Phase I focuses on analyzing historical data to determine whether the process is in control by estimating the in-control parameters of the process and the control limits. While in Phase II, the centre of attention is on monitoring on-line data to quickly detect shifts in the process from the in-control parameter values estimated in Phase I. Unusual observations in Phase I can lead to the inflation of control limits and reduction of power to detect process changes in Phase II. Therefore a successful Phase II analysis depends on a successful Phase I analysis in estimating in-control mean, variance, and covariance parameters. The preliminary data set collected in retrospective analysis involves either initial subgroups or individual observations. (Yahaya, et al., 2011).

- **Multivariate EWMA Control Models**

The scheme of the exponentially weighted moving average model the (EWMA) model was introduced by Roberts (1959). Crowder (1987) and Lucas and Saccucci (1990) provided excellent discussions on the EWMA model (Khoo&Teh,2009).Shewhart's control models have been the traditional tools for detecting larger shifts in the process mean (1.5 σ or more). For the univariate case, the EWMA is more effective than Shewhart control models in detecting smaller shifts in the process mean. When (n) measurements from each item are required, these univariate control models ignore the dependency among the (p) variables (Khaldi, 2007).

The multivariate exponentially weighted moving average control model (MEWMA) is the natural multivariate extension of the EWMA model proposed by Roberts (1959). It was introduced by lowry et al. (1992) and is more sensible in detecting nonrandom changes in the process and based on the principle of the weighted average of the previously observed vectors.

Despite the fact that it is used mainly for individual observation (n =1) it can be utilized in rational subgroup case. It is also a model for Phase II.

The MEWMA model has the statistics (Santos, 2012):

$$T_i^2 = Z_i' \Sigma_{Z_i}^{-1} Z_i > h \quad , \quad i = 1, 2, \ldots\ldots$$

$$Z_i = \lambda x_i + (1 - \lambda) Z_{i-1}$$

(1)

Where

(2)

where $Z_0 = 0$, λ is diagonal $p \times p$ matrix of the smoothing constant with $0 < \lambda_i \leq 1$, although in practice there is no reason to employ different values of λ in the same problem and x_i is the *i*th observation.

Lowry et al. (1992) provide two alternatives to compute the $\sum_z$, the exact covariance matrix:

$$\Sigma_{z_i} = \frac{\lambda[1-(1-\lambda)^{2i}]}{2-\lambda}(\Sigma) \qquad (3)$$

and the named asymptotic covariance matrix

$$\Sigma_{z_i} = \frac{\lambda}{2-\lambda}(\Sigma) \qquad (4)$$

Moreover, they point out that the ARL performance of the model depends only on noncentrality parameter θ:

$$\theta = [(\mu_1 - \mu_0)'\Sigma(\mu_1 - \mu_0)]^{1/2} \qquad (5)$$

where μ_1 is the mean vector for phase II. Notice that when $\lambda = 1$ MEWMA model is transformed on T^2 model.

One of the main troubles on this model is the selection of the h or UCL. Parbhu and Runger (1997) presented computed tables, base on the Markov chain approach, to choose the UCL according to the parameters λ, p, θ and ARL.

One the other hand, Bodden and Rigdon (1999) proposed a FORTRAN program to compute either the UCL for given values of ARL, λ, and p or ARL values of UCL, λ, and p.

Prabhu and Runger (1997) have provided a thorough analysis of the average run length performance of the MEWMA control model, using a modification of the Brook and Evans (1972) Markov chain approach. They give tables and models to guide selection of the upper control limit—say, UCL = *H*—for the MEWMA.

- **Multivariate CUSUM Control Models**

The Cumulative Sum (CUSUM) model was first developed by Page (1954) to detect slight but sustained shifts in the process level (1.5σ or less).

The CUSUM model is constructed for monitoring the mean of a process. It can be constructed for both individual observations n = 1 and the averages of rational subgroups n > 1 (Johnson, 1994).

Following the univariate design first introduced by Page (1954), multiple CUSUM models being used to monitor multiple variables was common practice (Woodall & Ncube, 1985). Rather than working with multiple CUSUM models,

Woodall and Ncube (1985) suggested creating a single control model to monitor multiple variables, called the Multivariate Cumulative Summation (MCUSUM) control model, the MCUSUM control model appears as the multivariate extension of the CUSUM control model originally proposed by page (1961).

It is focused on improving the sensitivity regarding the previously introduced T^2 model by detecting small shifts on the process and is based on the principle of accumulating information of the former observations. As well as the MEWMA model, MCUSUM is a Phase II model.

There are four main alternatives accepted to construct an MCUSUM model which is exposed below.

The first of these suggestions was introduced by Woodall and Ncube (1985). They proposed the individual monitoring of the mean vector through the utilization of univariate CUSUM models. Analogous to CUSUM there is also two-side model.

Its statistic is given by (Santos, 2012):

$$S_{i,j}^{-} = \min\left\{\begin{matrix} 0 \\ S_{i-1,j}^{-} + \dfrac{\bar{X}_{i,j} - \mu_{0,j}}{\sigma_{0,j} / \sqrt{n}} + k_j^{-} \end{matrix}\right\}$$

(6)

$$S_{i,j}^{+} = \max\left\{\begin{matrix} 0 \\ S_{i-1,j}^{+} + \dfrac{\bar{X}_{i,j} - \mu_{0,j}}{\sigma_{0,j} / \sqrt{n}} + k_j^{+} \end{matrix}\right\},$$

where $\mu_{0,j}$ is the j*th* element of the μ vector, $\sigma_{0,j}$ is the $(j\times j)$*th* diagonal element of Σ matrix, and k is a constant. Notice that when i =1 then $S_{i,j}^{-}$ and $S_{i,j}^{+} = 0$.

After that, Healy (1987) suggested a procedure to detect shifts in mean based on the linear combination of variables:

$$S_i = \max\begin{Bmatrix} 0 \\ S_{i-1} + a'\bar{X}_i - k \end{Bmatrix}, \tag{7}$$

Where

$$a' = \frac{(\mu_1 - \mu_0)'\left(\frac{\Sigma_0}{n}\right)^{-1}}{\left[(\mu_1 - \mu_0)'\left(\frac{\Sigma_0}{n}\right)^{-1}(\mu_1 - \mu_o)\right]^{1/2}} \tag{8}$$

and

$$k = 0.5\,\frac{(\mu_1 - \mu_0)'\left(\frac{\Sigma_0}{n}\right)^{-1}(\mu_1 - \mu_0)}{\left[(\mu_1 - \mu_0)'\left(\frac{\Sigma_0}{n}\right)^{-1}(\mu_1 - \mu_0)\right]^{1/2}} \tag{9}$$

On the other hand, Crosier (1988) presented two multivariate procedures. Here we present the version of the better ARL performance.

The statistics is

$$T_i^2 = \left[S_i'\left(\frac{\Sigma}{n}\right)^{-1} S_i\right]^{1/2} > h, \tag{10}$$

where

$$S_i = \begin{cases} 0 & if\ C_i \le k \\ (S_{i-1} + \bar{X}_i - \mu_0)(1 - \frac{k}{c_i}) & if\ C_i > k \end{cases} \tag{11}$$

where $S_0 = 0, k > 0$, and

$$C_i = \left[(S_{i-1} + \bar{X}_i - \mu_0)'\left(\frac{\Sigma}{n}\right)^{-1}(S_{i-1} + \bar{X}_i - \mu_o)\right]^{1/2} \tag{12}$$

Finally Pignatiello and Runger (1990) proposed likewise two MCUSUM models, the following resulting as the better performance alternative:

$$T_i^2 = \max \begin{cases} 0 \\ \left[S_i' \left(\frac{\Sigma}{n} \right)^{-1} S_i \right]^{1/2} - k n_i \end{cases} \tag{13}$$

where

$$S_i = \sum_{j=i-n_i+1}^{i} (\overline{x}_i - \mu_0) \tag{14}$$

and
$$n_i = \begin{cases} n_{i-1} + 1 & if \quad T_{i-1}^2 > 0 \\ 1 & otherwise \end{cases} \tag{15}$$

5.3 Performance of Multivariate Control Models

The evaluation and comparison of different types of multivariate control models are performed using statistical and economic performance indicators. The average number of samples collected up to the appearance of an out-of control signal (ARL), is the most commonly used statistical indicator to evaluate the performance of a control model and to make comparisons between different types of models. The ARL is a parameter that takes into account the probabilities of Type I and Type II errors. Therefore, to evaluate the parameters of a control model, it is customary to study the behavior of the ARL. It is desirable that the ARL of the model is large when the process is under control and quite small when the process is out of control. Accurate determination of the ARL is not always possible because the majority of control variables involve correlation. However, there are numerical methods for determining parameters that optimize control models behavior such as the Integral Equation Method, Markov Chains, and Simulation.

In recent decades, great deal of research has been conducted on the improvement and application of numerical methods to obtain approximate parameters for evaluating the performance of the univariate control model. However, when it comes to optimizing the parameters of a multivariate control model, few studies have been developed, with the exception of the MCUSUM control model. Lowry *et al.* proposed a table for the MEWMA with ARL k and h using the Simulation Method for an under control ARL of 200 and $p = 2$; 3; and 4 quality characteristics. Lee and Khoo applied the Markov Chain Method for situations under control with parameters ARL, k and h for the MCUSUM model for individual observations with $p = 2; 3;$ and 4 quality characteristics under control for ARL of 100, 200, 370, 500, and 1000. The Integral Equation Method with Gaussian Quadrature was proposed by Alves to optimize the ARL, k, and h in the MCUSUM control model for individual observations with $p = 2; 3;$ and 4 for quality characteristics under control for ARL of 200, 500, and 1000. This method involves the analytical derivation of an integral equation, whose numerical solution via Gaussian Quadrature enables the user to obtain the approximation solution of these parameters. This method is an excellent alternative for the optimization of the MCUSUM model, and since it is more versatile, it provides better results for the value of the ARL and a faster calculation method compared to simulations and relative simplicity of implementation. (Alves et al., 2013)

6. Data Analysis & Discussion Results

Control models are widely used as process monitoring tools, primarily to detect changes in the process mean or in its standard deviation, which can indicate deterioration in quality. Quality control problems arise when processes or products with two or more related quality variables are to be monitored or controlled such as water quality, concrete blocks and the work in hospitals ...etc. Multivariate Statistical Process Control consists of a number of powerful tools for problem solving and improvement of quality control by reducing variability in industrial manufacturing processes. Among these tools, the most commonly used statistical methods in industries are the multivariate control models (Alves et. al, 2013).

Water quality is one of the most important factors that must be considered when evaluating the sustainable development of a given region. (Cordoba et al., 2010). Water quality must be defined based on a set of physical and chemical variables that are closely related to the water's intended use. For each variable, acceptable and unacceptable values must then be defined. Water quality is considered the main factor controlling health and the state of disease.

The quality control is defined and applied in several fields, as in industries as in render services. In this study, the use of control model univariate and multivariate was fundamental to find the possible variables out of control in the drinking water in Khan-Younis Governorate. The analyzed variables are: Nitrate, Chloride and Total dissolved of solids.

6.1 Data and variables description

As most other Statistical quality control studies, data are difficult to get. For the purpose of the application of the methods in practical situation in the Khan-Younis Governorate we get a rich database for 52 samples .The samples taken from the groundwater wells at two stages, the first stage is in the autumn and the second stage is in the spring. Thus we have a database on 3 variables for the samples of drinking water.

The data come from the health laboratories for nutrition and water that follows to the Palestinian Ministry of Health from 1987 to 2012, we take three characteristic for the chemical component of water to monitor the quality of water which are:

Chloride (CL), Nitrate (NO3) & Total Dissolved of Solids (TDS) which are measurement of inorganic salts, organic matter.

6.2 Normality of Variables

Normality by Kolmogorov-Smirnov test: (Table 2) shows the normality of the three variable because p-value is higher than 0.05.

(Table 2): Kolmogorov-Smirnov test for the chemical water components

Variable	P-value
Chloride	0.644
TDS	0.780
Nitrate	0.813

6.3 Methods of data analysis

The data were labeled and recoded using the Spss V.20 and R 3.0.2 statistical software package and using R commander, from "Multivariate Statistical quality Control" package, we use the "MSQC" to create a Hotelling, MCUSUM & MEWMA, and from "Quality Control Models" package, we use "qcc" to create Xbar, EWMA and CUSUM models.

There are three models that we use to monitor the quality of water:

1). The first model displays the Shewhart Xbar model for the data set.

2). The second model displays a univariate control models such as CUSUM model for each using specific *h* and *k* values as determined and EWMA Control model for the data set and the detection of small shifts to monitor the quality of water.

comparing between them to decide which one of them is the best in the detection of small shifts to monitor the quality of water.

3). The third model displays a multivariate control models such as Hotelling T^2 model for the data set, next we want to displays a MEWMA model for the same data set.

4). The fourth model displays a MCUSUM control model using specific h and k values.

5). Finally we will compare between them to know which one of them is the best in the detection of small shifts .

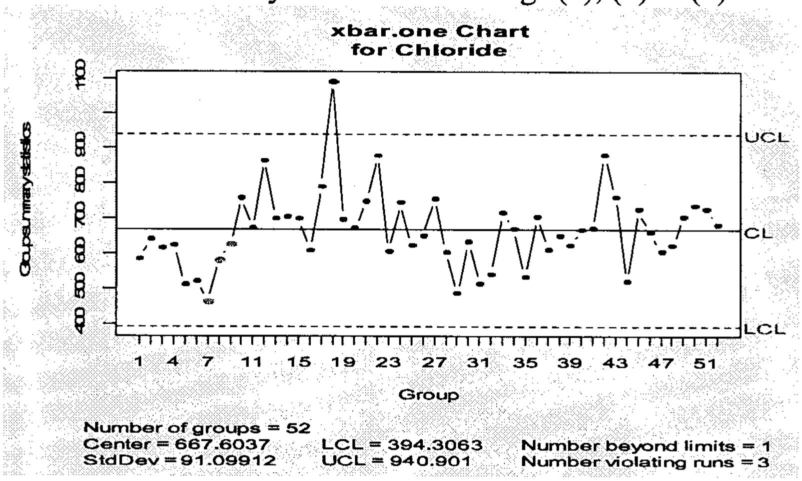
Then, we will compare between univariate & multivariate control models and between the multivariate control models to decide which the best model.

6.4 Analysis of Data:

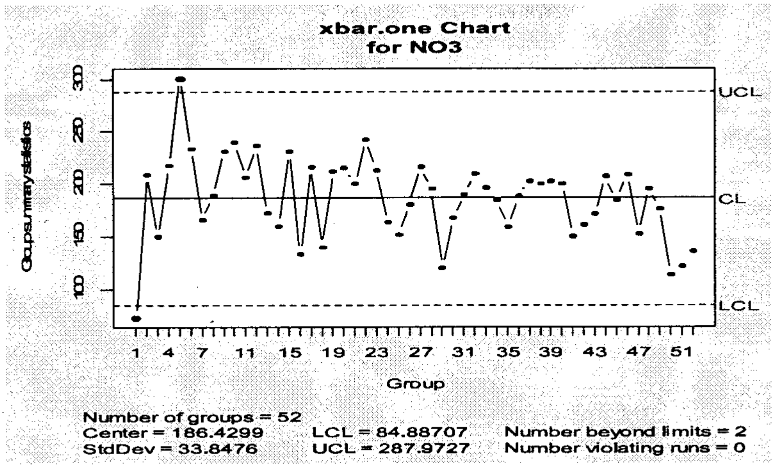
First we plot shewhart Xbar model for three variables (CL, NO3& TDS) to monitor the quality of water.

6.4.1 Results of the Analysis of Shewhart Xbar for the Chemical Water Components:

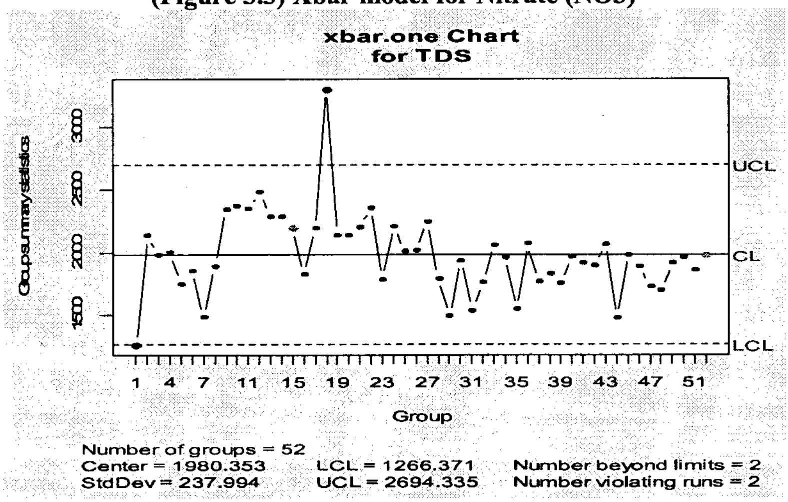
Plots were prepared for each of the three chemical component of water and the results analyzed as follows in figs (1), (2) & (3).



(Figure 3.1) Xbar model for Chloride



(Figure 3.3) Xbar model for Nitrate (NO3)



(Figure 3.2) Xbar model for TDS

As long as the readings remain randomly within the range between the LCL and UCL, the process is considered within control. But

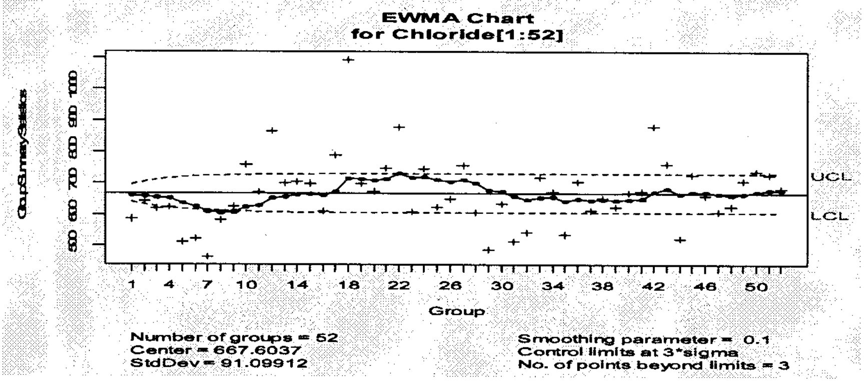
Figure (1) observed that only one sample goes beyond the upper control limit of chloride.

Figure (2) observed that only two samples go beyond the upper control and lower control limits of NO3.

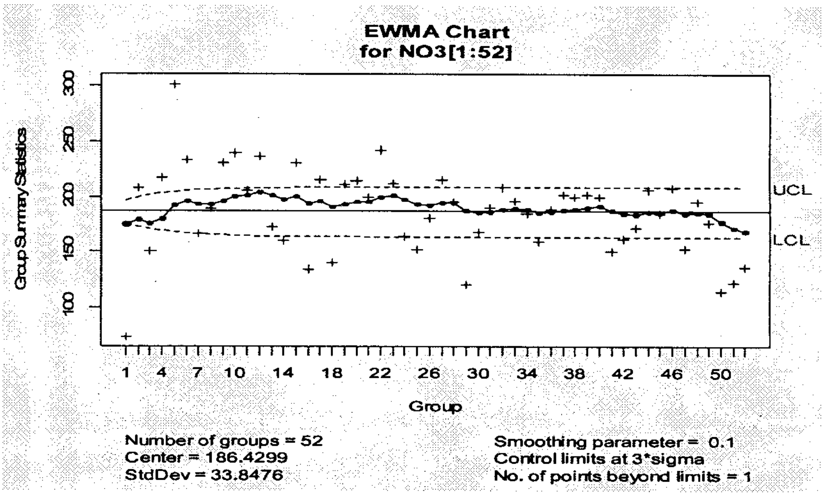
Figure (3) observed that only two samples go beyond the upper and lower control limits of TDS.

6.4.2 Results of the Analysis of EWMA Plot for the Chemical Water Components

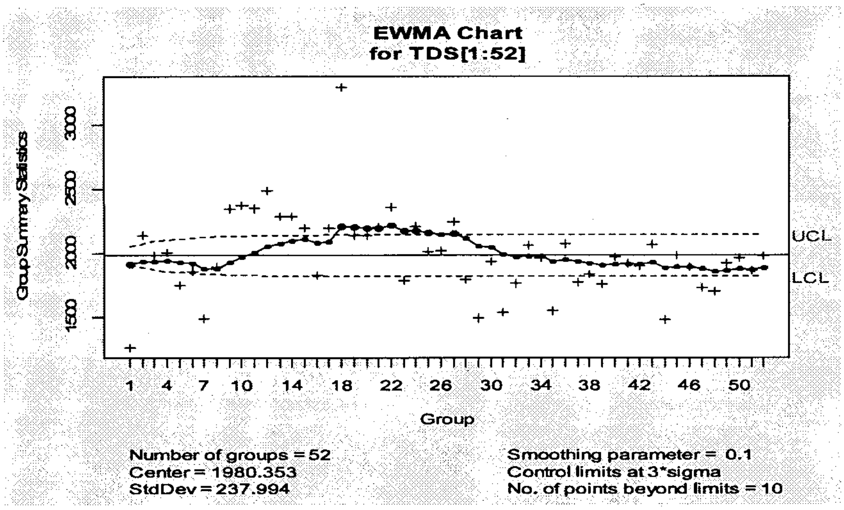
Plots were prepared for each of the three chemical component of water and the results analyzed as follows.



(Fig. 4): Conventional EWMA Plot for Chloride



(Fig. 5) Conventional EWMA Plot for NO3



(Fig. 6) Conventional EMMA Plot for TDS

We choose weighting constant (λ = 0.1) typically for the standard specification for quality of drinking water.

Analysis of the Graphical Trend Obtained in Figures of EWMA for (CL, TDS & NO3)

Figure (4) shows a high fluctuation of decreasing from sample No. 1 to 8 and increasing form sample No. 9 to 19 and from sample 20 upto the last sample and shows that the Chloride (CL) falls below the standard Specification of *250 mg/l.*

The trend of the graph for the conventional EWMA plot shows that the sample Nos. 7, 8, and 9 exceed the control limits; this means that the process is considered to be out of control.

Figure (5) shows a fluctuation of decreasing and decreasing in the NO3 component sample and shows that the NO3 falls below the standard Specification of *50 mg/l.*

Figure (6) shows a high fluctuation of increasing and decreasing in the TDS component sample and shows that the TDS falls below the standard Specification of *1500 mg/l.*

By observing the properties of chemical water, it's clearly that the process in identical to the standard specification.

6.3.3 Results of the Analysis of CUSUM Plot for the Chemical Water Components

Plots were prepared for each of the three chemical water components and the results analyzed as follows:

When k is selected to be 1, the parameter h is usually set at values of 4 or 5. The parameter h is the value against which the cumulative sum in the CUSUM scheme will be compared. In the context of groundwater monitoring, a value of h = 5 is recommended (Starks, 1988; Lucas, 1982).

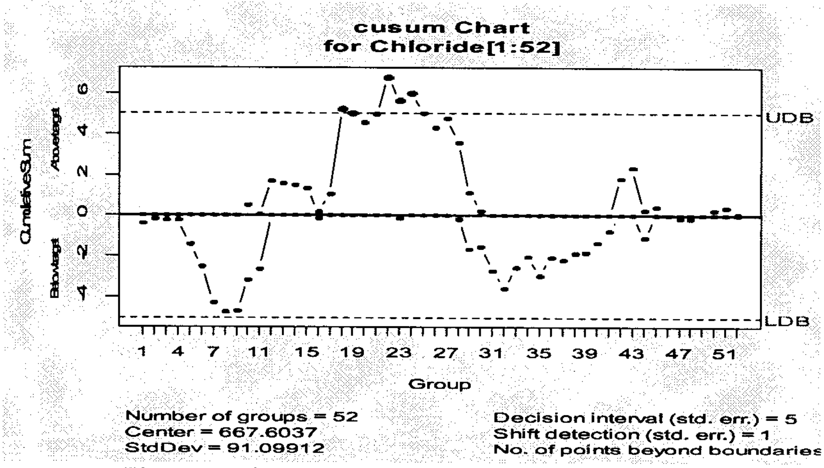


Figure (7) Conventional CUSUM plot for Chloride

Analysis of the Graphical Trend Obtained in Figure (7):

1) In generally the CUSUM plot shows a negative trend, acute fluctuation decrease and increase in the Chloride trend.
2) The trend of the graph for the conventional CUSUM plot shows that the Chloride non-conforming to the standard specification.

The out of control process and high degree of fluctuation in the CUSUM plot means that the quality of water is out of control.

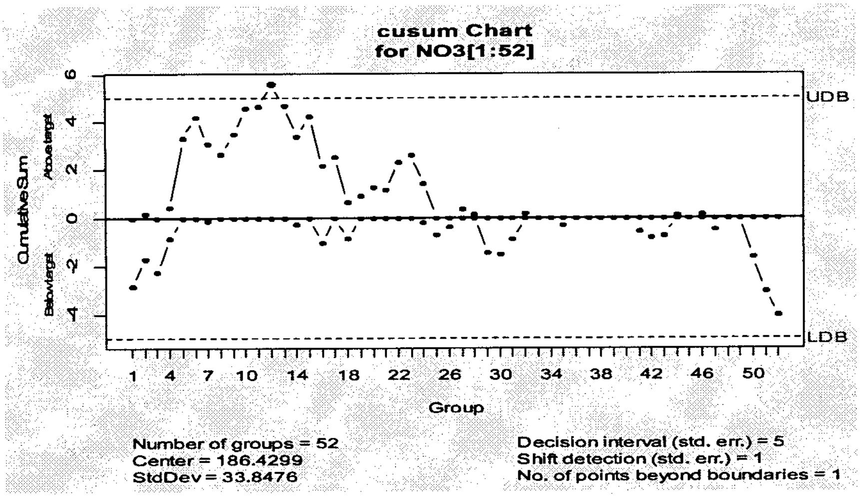


Figure (8) Conventional CUUSM plot for NO3

Figure (8) for Conventional CUSUM plot shows a fluctuation in decrease and increase in the NO3 trend , the trend non-conforming to the standard specification of NO3 and one Sample exceed the decision interval H^+, this means that the process is considered to be out of control.

The out of control process and high degree of fluctuation in the CUSUM plot means that the quality of water is out of control.

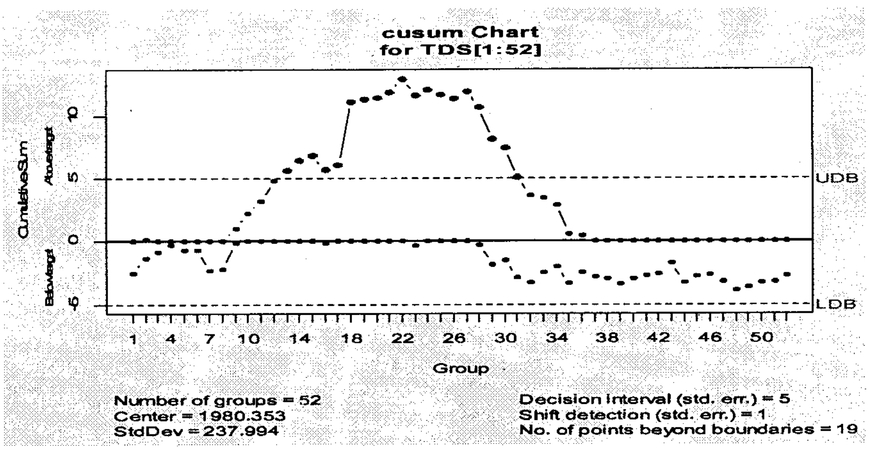


Figure (9) Conventional CUSUM plot for TDS

Analysis of the Graphical Trend Obtained in Figure (9):

- The trend of the graph for the conventional CUSUM plot shows a hard increase and decreasing in the TDS trend.
- The trend of the graph for the conventional CUSUM plot shows that the TDS.

In the next we will analyze the chemical component of water using multivariate control models.

6.4.4 Results of the Analysis of Hotelling Phase I Plot for the Chemical Water Components

Phase I focuses on analyzing data to determine whether the process is in control by estimating the in-control parameters of the process and the control limits.

Now, we can plot the Hotelling Model and from it we can decide which the process is in control or out -of-control.

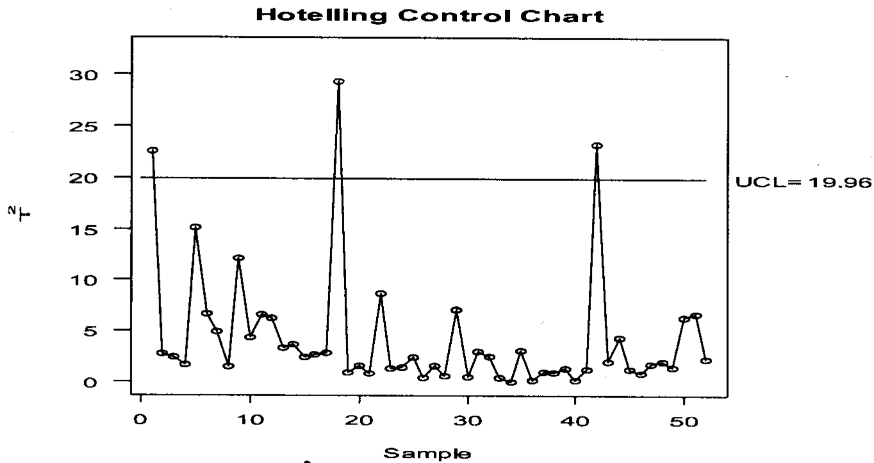


Figure (10) T^2 Hotelling control model in Phase I

Figure (10) show that there is fluctuation between increases and decreases in the three chemical component of water and the sample Nos. 1 , 18 & 42 fall outside of the upper control limits which is mean that the chemical component of water non-conforming to the standard specification of the quality of drinking water, so the process is out - of- control.

Note: In Phase II, the center of attention is on monitoring on-line data to quickly detect shifts in the process from the in-control parameters values estimated in Phase I.

6.4.5 Analysis of Multivariate EWMA Plot for the Chemical water Components

The multivariate exponentially weighted moving average (MEWMA) is another type of multivariate control models to monitor the quality of chemical.

Then, we can plot the Multivariate EWMA control model for the chemical water components as observed in Figure (11).

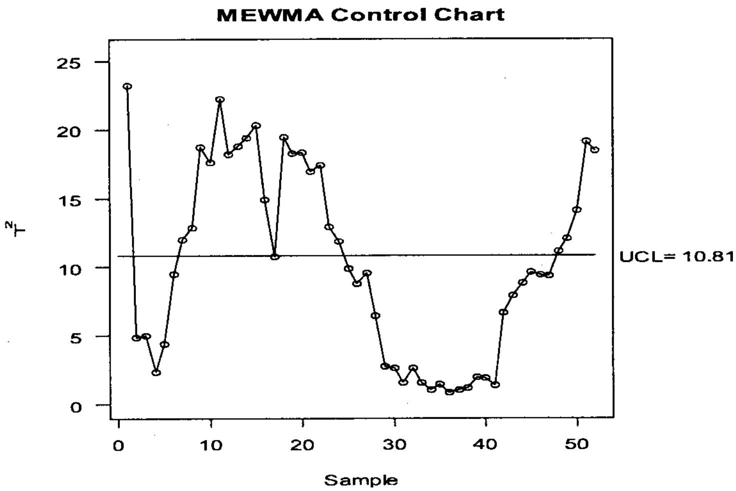


Figure (11) Multivariate EWMA for the chemical water components

the plot in figure (11) tells us that the process is out of control because some of MEWMAt lie out of the upper control limit. However, there seems to be that,

- There is decrease in sample No. 1 to sample No. 4 , and fluctuation between increase and decrease from sample No 11 upto sample No. 24
- A high decrease from sample No. 18 upto sample No. 38.
- A high increase from sample No. 5 upto sample No. 15 and from sample No. 39 upto the last sample.
- The trend of the graph show that there is a sample lies outside of the upper control limit which is , sample No. 1 , from sample No. 7 upto sample No. 16 , sample No. 18 upto sample No. 24 and sample No. 48 upto sample No. 52, this means that the process is considered to be out of control.

A fluctuation of the trend of the graph between increase and decrease reveals the fact that the drinking water in Khan-younis governorate is non- conforming to the standard specifications and not suitable for human use.

6.4.6 Results Analysis of Multivariate CUSUM Plot for the Chemical Water Components

The multivariate CUSUM is the third type of multivariate control models to monitor the quality of the chemical water components and there are two type of MCUSUM (Crosier 1988 & Pignatiello 1990).

A) MCUSUM by Crosier 1988:

we can plot the Multivariate CUSUM control model (Crosier 1988) for the chemical water components as observed in Figure (12).

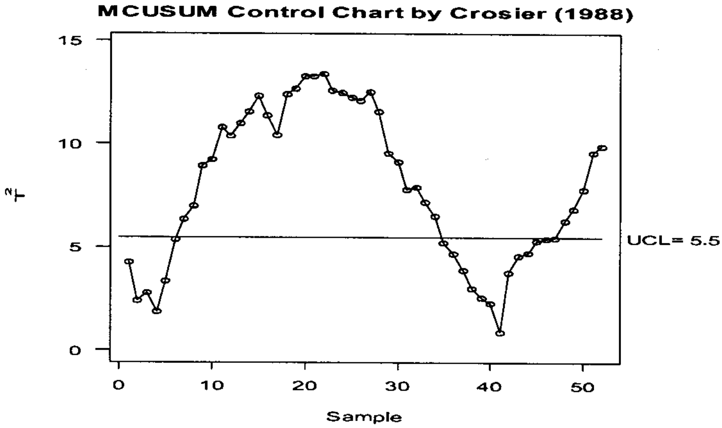


Figure (12) Multivariate CUSUM by Crosier (1988) for the chemical water components

the plot in figure (12) tells us that the process is out of control because some of MCUSUM lie out of the upper control limit. However, there seems to be that,

- There is an high increase in sample No. 5 upto sample No. 22 and from sample No.42 upto sample No. 52
- There is a decrease from sample No. upto sample No. 4 & there is a high decrease from sample No. 23 upto sample No. 41
- The trend of the graph show that there are a sample lie outside of the upper control limit which is , sample No. 7 upto sample No. 34 and from sample No. 48 upto sample No. 52, this means that the process is considered to be out of control.

B) MCUSUM by Pignatiello (1990):

we can plot the Multivariate CUSUM control model for the chemical water components as observed in Figure(13)

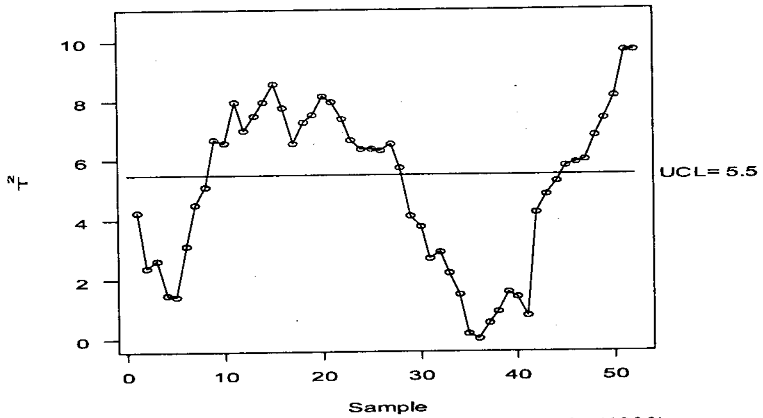


Figure (13) Multivariate CUSUM by Pignatiello (1990) for the chemical water components

the plot in figure (13) tells us that the process is out of control because some of MCUSUM lie out of the upper control limit. However, there seems to be that,

- ★ There is increase from sample No. 6 upto sample No. 1 and a high increase from sample No.37 upto the last sample.
- ★ There is decrease from sample No. 1 upto sample No.5 and a high decrease from sample No. 28 upto sample No. 36.
- ★ There is a fluctuation between increase and decrease from sample No. 9 upto sample No. 27.
- ★ The trend of the graph show that there are a sample lie outside of the upper control limit which is , sample No. 9 upto sample No. 28 and sample No. 45 upto the last sample, this means that the process is considered to be out of control.

7. Conclusion

Maintaining quality of drinking water satisfying the standard specification is a very important vital requirement, in this study we plot some conclusions a concerning using multivariate control models such as (Hotelling's, MEWMA, MCUSUM) to monitor the quality of chemical component of drinking water, so we reached to the following conclusion:

1) A high degree of fluctuation between increase and decrease to the quality of drinking water and the samples that lies outside of the upper or the lower control limits means that the process is out-of-control and nonconforming to the standard specifications.

2) **In the Univariate case :**

A) The CUSUM and EWMA are better than the Shewhart models.
B) EWMA models tended to be slightly slower in providing significant, but were much easier.
C) The CUSUM model is faster in detecting the out of control.
D) The CUSUM model is more effective in detecting the small shifts than the EWMA model because it had a decision interval (h) and reference value (K).

3) **In Multivariate case :**

E) The MEWMA is more effective than univariate EWMA in detecting small shifts.
F) The MCUSUM is more effective than univariate CUSUM in detecting small shifts.
G) The Hotelling's control model is better than the Shewhart models in monitoring the quality of drinking water.
H) The MEWMA is more effective than Hotelling's control model in detecting the variations in the process.
I) The MCUSUM model is more effective than the MEWMA and the Hotelling's control models in detecting the small shifts in monitor the quality in drinking water.
J) MCUSUM in Crosier (1988) is better than MCUSUM in Pignatiello (1990) for detecting the small shifts and variations in monitor the quality of drinking water.

We can conclude finally that the MCUSUM in Crosier (1988) is the best model to monitor the quality of drinking water.

8. Recommendations

The following recommendations for future studies are suggested:

1. This research work can be extended by applying the Multivariate control models such as MCUSUM & MEWMA models in other sectors of construction industry like food

products factories by monitoring the basic components of these materials.

2. The use of Multivariate control models are not confined to the chemical component of water (CL, NO3 & TDS) it can be extended to monitor other components of water such as calcium, sodium & Phosphate.
3. Using Multivariate quality Control Models to monitor the quality of the Palestinian hospitals work.
4. Detecting and interpretation for out-of-control signals in multivariate statistical quality control.

References

[1] Alt, F. B. (1985) "Multivariate Quality Control", in S. K. a. N. L. Johnson (eds) "Encyclopaedia of Statistical Sciences", Vol. 110-122, John Wiley, New York.

[2] Alves C., Henning E., Samohyl R., Konrath A. and Walter O.(2013), "Application of Multivariate Control charts For Monitoring An Process" , vol. 17, no. 2, TECNO-LÓGICA, Santa Cruz do Industrial Sul.

[3] Brook, D., & Evans, D. A. (1972). An Approach to the Probability distribution of CUSUM Run Length. Biometrika, 5(3), 539-549.

[4] Champ, C. W. and Woodall, W. H. (1987), "Exact Results for Shewhart Control Charts with Supplementary Runs Rules". Technometrics 29, 4, 393-399

[5] Chanda, M. J. (2001). Statistical Quality Control. CRC Press, LLC, 2000 N.W. Corporate Blvd., Boca Raton, Florida 33431.

[6] Croiser, R. B. (1988). Multivariate Generations of Cumulative Sum Quality control Schemes Technometrics Vol. 30, 219-303.

[7] Crowder, S.V.(1987),"A simple method for studying run length distributions for exponentially weighted moving average control charts" Technometrics, vol.29, pp.401-407.

[8] Dai Y. , Zhou C. & Wang Z. (2009)," Multivariate Cusum Control Charts Based on data depth For Preliminary Analysis", Tech. Report , Department of Statistics, Nankai University, Tianjin 300071, PR, China.

[9] Dou, Y. and Ping, S. 2002. "One Sided Control Charts for The Mean of Positively Skewed Distributions," The Quality Management. Vol. 13, pp. 1021-1033.

[10] Hawkins, D. M. and Olwell, D. H. (1998). Cumulative sum charts and charting for quality improvement New York :Springer,1998.

[11] Healy, J. D. (1987). A Note on Multivariate CUSUM Procedures. Technometrics,vol.29,No.4, pp 409-412

[12] Hotelling H.(1947). Multivariate Quality Control., illustrated by the air testing of sample bombsights. Techniques of Statistical Analysis, 111-184. New York, McGraw Hill.

[13] Johnson, R.A., (1994), Miller & Freund's Probability & Statistics For Engineers, Fifth Edition, Prentice Hall, Englewood Cliffs, New Jersey.

[14] Khalidi, Mohammed (2007)," Multivariate Quality Control : Statistical Performance And Economic Feasibility " , MSc Dissertation, Wichita State University.

[15] Kolarik, W. J. (1999). Creating Quality: Process Design for Results. McGraw-Hill International Edition. Singapore, McGraw-Hill Book Company.

[16] Kosalprapai, Anawat (2002)," Multivariate Attribute P Chart For Both Correlated and Uncorrelated Quality Characteristics" , MSc Dissertation, Kasetsart University.

[17] Lowry, C. A. ; Woodall, W. H. ; Champ, C. W. and Rigdon, S. E. (1992) A Multivariate Exponentially Weighted Moving Average Control Chart Technometrics Vol. 34, No. 1,pp.46-53.

[18] Lucas, J. M. 1982. "Combined Shewhart-CUSUM Quality Control Schemes." Journal of Quality Technology. Vol. 14, pp. 51-59.

[19] Lucas, James M.; Saccucci, Michael S (1990). Exponentially weighted moving average control schemes: properties and enhancements.With discussion and a reply by the authors Technometrics,Vol. 32, no. 1, pp.1–29.

[20] Mason, R. L., & Young, J. C. (2002). Multivariate Statistical Process Control with Industrial Applications. Philadelphia: ASA-SIAM.

[21] Montgomery, D. C. (1996). Introduction to Statistical Quality Control 3rd. ed. Jon Wiley and Sons, New York, NY.

[22] Page (1954). Continuous Inspection Schemes Biometrika Vol. 41, 100-114.

[23] Pignatiello, P.P. (1990)," Comparison of Multivariate CUSUM Charts ", Journal of Quality Technology, vol.22, No.3, pp.173-186

[24] Prabhu, S.S., Runger, G.C. (1997), "Designing a multivariate EWMA control chart. Journal of Quality Technology" ,vol.29, No.1,pp.8-15.

[25] Roberts, S.W. (1959), Control Charts Based on Geometric Moving Averages, Technometrics, vol.1, Issue 3, pp. 239-250

[26] Runger, G.C. & Montgomery, D.C. (1997) Multivariate and univariate process control: geometry and shift directions. Quality and Reliability Engineering International, vol.13, Issue 3, pp.153-158.

[27] Santos, Edger F.(2012)," Multivariate Statistical Quality Control Using R ",vol.14, Springer, New York, pp.17-19 , 46-52.

[28] Woodall, W. H. and Montgomery, D. C. (1999). Research issues and ideas in statistical process control. Journal of Quality Technology, 31(4), 376-386.

[29] Woodall. W. H and Ncube. M. M. (1985). Multivariate CUSUM Quality Control Process Technometrics Vol. 27, No. 3, 285-292.

[30] Yahaya S. , Ali H. & Omar Z. (2011), " An Alternative Hotelling T^2 Control Variance (MVV) Chart Based on Minimum Vector.

[31] Zhao Tao,(2007),"A New Method for Detection and Classification of Out- of- control Signals in Autocorrected Multivariate Processes " pp.1-2