The Scientific Journal of Cihan University – Sulaimanyia PP: 109-129
Volume (3), Issue (1), Jun 2019
ISSN 2520-7377 (Online), ISSN 2520-5102 (Print)

Constructing Control Charts, for Monitoring Chemical Parameters of Water in Kanisard Factory in Sulaimani City

Assist.Prof. Dr. Kawa M. Jamal Rashid [1], Gulshan Star Muhammad[2]

[1,2]Statistics and Informatics Department, College of Administration and Economics, Sulaimani University, Sulaimani, Kurdistan Region of Iraq
kawa.rashid@univsul.edu.iq[1]
gulshan.star@yahoo.com[2]

Abstract

Monitoring of the production process is an important subject for developing the quality of the product and reducing the costs, (ARIMA) residual chart is a special control chart used to specify and detect the quality behavior in time-correlated process data, to determine if they are out of control or in control. Furthermore, this type of chart is useful for adjusting and specifying the quality limits during the process. Water quality is considered the main factor of controlling human health in disease therefore, it is necessary to keep the quality of drinking water to be in control. The main objective of this study is to monitor the two important chemical parameters of drinking water which are Power of Hydrogen (PH) and Magnesium (MG), it also aims to determining the control limits for both (PH &MG) from the optimal tolerance limits to control the water production for reach better quality products in the future, by taking the data for each of the parameters (PH& MG) from January to September (2018) from the (KANISARD) factory for producing drinking water at Sulaimani city in Kurdistan region in Iraq. By using the autoregressive integrated moving average (ARIMA) control chart. The result of the study showed that both (PH &MG) processes are in statistical control by using ARIMA control chart, and also the optimal tolerance limits were determined for both of the parameters.

Keyword: *quality control, statistical process control, control chart, ARIMA control chart, tolerance limit, PH, MG.*

الملخص

يعد المراقبة عملية الإنتاج موضوعًا مهمًا لتطوير جودة المنتج وخفض التكاليف،ان لوحة السيطرة (الانحدارالذاتي والاوساط المتحركة المتكاملة) هي عبارة عن لوحات السيطرة الخاصة،التي تستخدم لتحديد واكتشاف سلوك الجودة في بيانات العملية المرتبطة بالوقت إذا كانت خارجة عن السيطرة أوتحت السيطرة، كما أن هذا نوع من اللوحة مفيد لضبط الحدود الجودة وتحديدها أثناء العملية. تعتبرنوعية المياه العامل الرئيسي في السيطرة على صحة الإنسان من الأمراض، لذلك من الضروري الحفاظ على جودة المياه الشرب والسيطرة عليها. الهدف الرئيسي من هذه الدراسة هو مراقبة المعلمتين الكيميائيتين المهمتين لمياه الشرب وهما قوة الهيدروجين والمكنسيوم، كما تهدف أيضًا الى تحديد الحدود التحكم لكل من قوة الهيدروجين والمكنسيوم من حدود التسامح المثالية للتحكم في إنتاج المياه من أجل الوصول إلى منتجات ذات جودة أفضل في المستقبل، من خلال أخذ البيانات لكل من المعلمات قوة الهيدروجين والمكنسيوم من ينايرإلى سبتمبر 2018من مصنع (كانيسارد) لإنتاج مياه الشرب في مدينة السليمانية في إقليم كردستان في العراق. باستخدام لوحة السيطرة (الانحدارالذاتي والاوساط المتحركة المتكاملة). أظهرت نتائج الدراسة أن كلا من قوة الهيدروجين والمكنسيوم تحت السيطرة الإحصائية وكذلك تم تحديد حدود التسامح المثلي لكلا المعلمتين.

The Scientific Journal of Cihan University – Sulaimanyia PP: 109-129
Volume (3), Issue (1), Jun 2019
ISSN 2520-7377 (Online), ISSN 2520-5102 (Print)

پوختە

جاودیریکردنی بروسه ی به رهه مهینان بابه تیکی کرنکه بو بیشکه وتنی به رهه م و که مکردنه وه ی تیجون، نه خشه ی کونترولی (ئه ریما) یه کیکه له نه خشه کونترولیه تاییه ته کان که به کارئه هینری بو دیاری کردن و که شفکردنی سلوکی جوربو ئه و داتایانه ی که به یوه ستن به کاته وه بوزانینی ئه وه ی که بروسه که له ده ره وه ی کونترولە یان له ناو کونترولە، هه روه ها ئه م جوره نه خشه ی کونترولە به سوده بوجاککردن و دیاری کردنی سنوری جور له کاتی بروسه که دا. جوری ئاویه کیکه له فاکته ره سه ره کیه کان بو کونترولکردنی ته ندروستی مروف له نه خوشی له به رئه وه کرنکه که کوالیتی ئاوی خواردنه وه له زیر کونترول بهیلینه وه. ئامانجی سه ره کی له م تویزینه وه یه جاودیریکردنی دوو بیکهاته ی کرنکی ئاوه که ئه وانیش (هیزی هایدروجینی ومه کنسیوم) ن، وه هه روه ها دیاری کردنی باشترین سنور(واته زور ترین وکه مترین سنور) بو ئه وه ی له داهاتودا بکه ین به کوالتیه کی باشتر، ئه مه ش به وه رکرتنی داتا بو هه ریه ك له هیزی هایدروجینی ومه کنسیوم له مانکی یه ك بو مانکی نو 2018 له کارکه ی (کانیسارد) له شاری سلیمانی له هه ریمی کوردستان له عیراق. به به کارهینانی نه خشه ی کونترولی (ئه ریما). وه له ئه نجامی ئه م تویزینه وه یه ده رکه وت که (هیزی هایدروجینی ومه کنسیومیش) له زیرکونترولی ئاماریدان وه هه روه ها باشترین سنوریش دیاری کران بوهه ردوو بیکهاته ی ئاوه که.

1.1 Introduction:

Quality control (QC) is a significant function in the factory that deals with inspecting the product before transporting the product to customers. Hence, quality is one of the most widely important customers deciding factor for choosing among the competing services and products (Ashour, 2014). Statistical quality control (SQC) is a method that uses statistical techniques to monitor and control the quality of the product by using the control charts as test tools that frequently used to monitor the manufacturing process (Salih, 2011). Statistical process control (SPC) is defined as a powerful collection of solving tools problem which is very helpful for improving capability and getting the stability of the process by reducing the variability. Statistical process control (SPC) is one of the most important technological improvements in twenty century because it is based on sound underlying principles; it is easy and can be applied to any processes (Montgomery, 2009). Control charts are the main and most important tools in statistical process control (SPC), to detect the assignable causes only, which should be removed by engineering actions or operator. Control charts include three horizontal lines and these are, Central Line, Upper Control Limit, and Lower Control Limit. Control charts can be divided into two types of control charts and these are variable control chart and attribute control chart (Magaji et al, 2015). The control charts process contains taking samples from a process and plotting the control statistic on time order, which is calculated from the sample, the chart is said to give a signal and the process is considered to be out of control if the statistic plots fall outside the predetermined control limits and then it should make an effort for removing the cause of changing of the process. Furthermore, a process is considered to be in a state of statistical control if it has only common cause variation. Thus, a cause of a signal of the control charts is either common cause or special cause, where a common cause signal is called a false alarm. When there are changes in the process because of a special cause, thus the best control chart gives a signal rapidly and have a low false alarm rate where just inherent variation is present (Xiao, P., 2013). Furthermore, any production process is in statistical control if three or less than three points out of (1000) points fall outside the control limits, and also the process is out of control if more than three points of (1000) point approximately falls outside the control limits (Hama Rasoul et al, 2019).

1.2 Literature review:

This section gives the background about using ARIMA control chart in different areas and also other control charts that used in the quality of water:

DOI: http://dx.doi.org/10.25098/3.1.8

Volume (3), Issue (1), Jun 2019
ISSN 2520-7377 (Online), ISSN 2520-5102 (Print)

Ahmad, (2006) compared ARIMA residual chart with traditional R-chart for monitoring the cigarettes production, the result of the study showed that ARIMA chart is better than the R-chart in detecting shifts and removing autocorrelation in the process. George et al. (2009) applied the Hotelling T-square control chart for the fault detection of drinking water treatment. The result of the study showed that the Hotelling T-square control chart can be applied effectively for the fault detection of drinking water treatment.

Kovářík and Klímek (2012) studied the ARIMA residual chart in the financial data, and illustrate how time series control charts are sensitive in detecting small shifts. Ashour (2014) compares between univariate control charts (Shewhart, EWMA, CUSUM) and multivariate control charts (Hotelling, MEWMA, MCUSUM) to monitor the quality of three parameters (chloride, nitrate, total dissolved salts) of drinking water in KhanYounis governorate in Gaza. The result of the study showed that the multivariate cumulative sum (MCUSUM) is the best control chart to detect small shifts to monitor the drinking water quality in KhanYounis governorate in Gaza.

Tasdemir and Kowalczuk (2014) proposed the methodology for monitoring a plant scale copper flotation process based on statistical quality control charts. It was concluded that the ARIMA residual chart is more suitable and capable of monitoring the process than the standard control charts when the data have autocorrelation problem. Bhasin et al. (2016) used Shewhart (X-bar) control chart to evaluating the quality of water of a tropical river in India, analysis of different parameters such as dissolved oxygen, turbidity, total alkalinity, total hardness, chloride, and calcium was performed. The result of the study showed that the X-bar control chart gives a clear illustrative of the pollution condition of the river, and the sample means values in control chart, demonstrating the poor quality of the water.

Tasdemir, (2017) investigated determining control limits for ash content of clean coarse coal that produced by heavy medium drum at the coal preparation plant. The result of the study showed when considering non-normality and auto-correlation for ARIMA residual chart, the number of the points that exceed the control limit is less than those obtained by control charts using original data. Salleh et al. (2018) compared the performance of the Box-Jenkins method by using ARIMA residual chart with geometric Brownian motion method for monitoring the autocorrelation process by using the data from the furnace temperature. The result of the study revealed that both methods are performed similarly and gave the same result for monitoring process control and model accuracy, but the result also revealed that geometric Brownian motion method is easier in comparing with Box-Jenkins approach.

2-Theoretical part:

2.1 Autoregressive integrated moving average (ARIMA) control chart:

Traditional Shewhart control chart fails if the data has a very low degree of autocorrelation, and in control, diagram failure shows a large number of points outside the control limits because in traditional Shewhart statistical process control (SPC) it is assumed that the measured data are not auto-correlated. In the continuous processes case, this phenomenon is not unique. In the discrete processes case, the autocorrelation of the data becomes an increasingly frequent phenomenon. The time series stochastic modeling and using autoregressive integrated moving average (ARIMA) model is one of the methods to overcome the autocorrelation problem of the data. Based on the method of Box-Jenkins, linear stochastic auto-regressive models (AR), moving average (MA) model, (ARMA) models, and autoregressive integrated moving average (ARIMA) models, is a realization of the time series of a stochastic process, and they have a characteristic form of the

Volume (3), Issue (1), Jun 2019
ISSN 2520-7377 (Online), ISSN 2520-5102 (Print)

autocorrelation function (ACF) and partial autocorrelation function (PACF). Both of the (ACF) and (PACF) are the vital tools for giving information about the stochastic processes, and they are used for describing the models of the time series. In practice, very often there is a non-stationary process, because of changing the mean value or changing variance overtime non-stationary can be present (Kovářík et al, 2015; Kovářík & Klímek, 2012; Rashid, 2016).

The original integrated process is an autoregressive integrated moving average process (ARIMA) of order (p, d, q), and (p) are the number of the autoregressive term, (d) is the number of the differences and (q) is the number of the moving average term. Based on finding an appropriate time series model also using the control chart for residuals ARIMA control chart is work. The ARIMA (p, d, q) model has the general form:

$$\Phi p(B).\Delta^{d}.x_t = \theta q(B)\,\varepsilon_t \tag{1}$$

Where,

$\Phi p(B) = (1 - \Phi_1 B - \Phi_2 B^2 \ldots - \Phi_p B^{p})$ is autoregressive polynomials in (B) of p-th order

$\theta q(B) = (1 - \theta_1 B - \theta_2 B^2 \ldots - \theta_q B^{q})$ Is moving average polynomials in (B) of q-th order

(B): backshift operator where $B.x_t = x_{t-1}$

(Δ^{d}): $(1 - B)^{d}$ Is the suitable integer number of differencing in order to get a stationary time series and (t) is time.

$\Phi_1, \Phi_{2\ldots}\Phi_p$: are the autoregressive parameter models

$\theta_1, \theta_{2,} \ldots \theta_q$: are the moving average parameter models

ε_t: White noise (unpredictable and its values are uncorrelated, also it is distributed normally with zero mean and variance constant)

In practical application ARIMA model is the most widely used model, suppose the model is:

$$X_t = \xi + \Phi x_{t-1} + \varepsilon_t \tag{2}$$

Where ξ a ϕ $(-1 < \Phi < 1)$ are unknown constants, and ε_t is uncorrelated variable and distributed normal with zero mean and constant stander deviation, this equation is called first-order autoregressive model AR(1). If equation (2) expanded in the form

$$X_t = \xi + \Phi_1 x_{t-1} + \Phi_2 x_{t-2} + \varepsilon_t \tag{3}$$

The second order autoregressive model AR (2) is obtained. In general, the variable(x_t) is dependent in the previous value $x_{t-1}, x_{t-2}\ldots etc$, where the AR (p) has the form$[X_t = \xi + \Phi_1 x_{t-1} + \Phi_2 x_{t-2} + \cdots + \Phi_p x_{t-p} + \varepsilon_t]$. By using the random component (ε_t) if the data dependence is modeled, the MA (q) model is obtained. The moving average model of first-order MA (1) has a form:

$$X_t = \mu + \varepsilon_t - \theta\varepsilon_{t-1} \tag{4}$$

And MA (2) has the form:

$$X_t = \mu + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} \tag{5}$$

In general moving average of order (q) MA (q) have the form:

$$X_t = \mu + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} \ldots - \theta_q\varepsilon_{t-q} \tag{6}$$

The correlation between$(X_t and\, X_{t-1})$ have the form$\rho 1 = -\theta/(1 + \theta^2)$, this corresponds to the (ACF) shape. Frequently it is appropriate for modeling a mix of both autoregressive and moving

DOI: http://dx.doi.org/10.25098/3.1.8

Volume (3), Issue (1), Jun 2019
ISSN 2520-7377 (Online), ISSN 2520-5102 (Print)

averages model in modeling practical problems ARMA (p, q). ARMA of first-order model ARMA (1, 1) have the form

$$X_t = \xi + \Phi x_{t-1} + \varepsilon_t - \theta\varepsilon_{t-1} \tag{7}$$

The ARMA model supposes a stationary process that is meaning that the quality character reference value is around a constant mean. In practice, there are processes such as (chemical industry) where the monitored variable value is running away, in these situations it is appropriate for modeling the processes by using the suitable model with backward difference operator Δ , like ARIMA model (0,1,1) with the form

$$X_t = x_{t-1} + \varepsilon_t - \theta\varepsilon_{t-1} \tag{8}$$

Shewhart model is different from ARIMA models ($X_t = \mu + \varepsilon t$ for $t = 1,2, \dots.$) though, if ϕ=0 in equation $X_t = \xi + \Phi x_{t-1} + \varepsilon_t$ or $\theta = 0$ in the equation$X_t = \mu + \varepsilon_t - \theta\varepsilon_{t-1}$, the Shewhart model process is obtained. Selecting the suitable SPC control chart is another important step in using ARIMA models when residual testing determines that they are coming from the normal distributions and they are not auto-correlated (Kovářík et al, 2015; Kovářík & Klímek, 2012; Rashid, 2016). Furthermore, the original Box-Jenkins modeling procedure refers to the application of three steps process of model selection (identification, parameter estimation, and model checking) (Muhammad, 2011; Muhammad, 2018). First step is Identification: in this step data is plotted for the time series for checking if the data is stationary or not, if the data does not have any pattern such as seasonality or trend then the data is stationary, but if the data is not stationary then the transformation and difference are needed for making data stationary, and transformation of the data such as (logarithm, square root, etc) are helpful for stabilizing the variance, also (difference method) is helpful for stabilizing the mean, Box-Jenkins uses different graphs in addition, to the more application function defined as (ACF) and (PACF), in this step also the degree of ARIMA model (p, d, q) is determined for the data, by selecting this order that gives the minimum AKAIKE (MAIC) (Muhammad, 2011; Muhammad, 2018).

In addition, stationary can be determined by using the time series plot for the observation and also we can check the stationary for the time series by using the autocorrelation function, and this by taking the correlation coefficient value, and if the series is stationary then the value of the correlation coefficient approaching zero after lag two and three, but if the series is not stationary the value of the correlation coefficient approaching zero after a large number of lags may reach seven or eight and may not approach zero (Muhammad, 2011).AKAIKE information criterion (AIC): In (1973-1974) AKAIKE suggested AKAIKE information criterion (AIC), and this is a criterion that used to identify and select a suitable statistical model order, and popularly this criterion is used with ARIMA models to finding the suitable order of the model,

$$AIC = n \ln \sigma^2 a + 2m \tag{9}$$

Where,

n: is the number of observations

$\sigma^2 a$: White noise variance

m: is the number of the estimated parameters (Muhammad, 2011; Muhammad, 2018).

Table (1) gives the general rule to identify the models by using autocorrelation (ACF) and partial autocorrelation (PACF) plots to determine a proper model for the data (Muhammad, 2011; Muhammad, 2018):

Volume (3), Issue (1), Jun 2019
ISSN 2520-7377 (Online), ISSN 2520-5102 (Print)

Table (1) Model Identification

Model	ACF	PACF
AR(p)	Tail off	Cut off after (p)
MA(q)	Cut off after (q)	Tail off
ARMA (p,q)	Tail off	Tail off

The second step is Estimation process: estimation means finding the value of the model coefficients which provide the best fit to the data (Muhammad, 2011; Muhammad, 2018). The final step is Diagnostic model checking: this step is based on studying the autocorrelation plots for the estimated residuals(ε_t). To see how much the suggested model corresponds to the behavior of the time series (Muhammad, 2011; Muhammad, 2018).

First: Residual Autocorrelation test:

In testing the residual autocorrelation, if the autocorrelation coefficients of the residuals fall within the confidence limit, then that is mean that the residuals not systematic which r for residuals (a) is random and the identified model is an appropriate model and the confidence limits are:

$$-1.96\frac{1}{\sqrt{n}} \leq r_k(a_t) \leq 1.96\frac{1}{\sqrt{n}}$$

Where (n) is the number of observation, and (k) is the lag period (Muhammad, 2011; Hamarasoul et al, 2019).

Second: The goodness of fit test:

Box-pierce: In 1970 box and pierce suggested the statistic(Q) for the goodness of fit test for suggested ARMA model such that:

$$(Q) = n\sum_{k=1}^{h} r^2{}_k(\hat{a}_t) \sim x^2(h-m) \tag{10}$$

$H_0: p_k(a) = 0$ V/s $H_1: p_k(a) \neq 0$ Where,

(Q) Statistic: have the chi-square x^2 with $(h-m)$ degree of freedom

(n): Number of observation

(r_k): sample autocorrelation function at lag (k) of an appropriate time series$(\hat{a}_t)$

(h): is the largest lag used

(m): Number of estimated parameters of the identified model

And (Q) statistics compares with the $x^2 chi-square\ table\ with(h-m)$ degree of freedom with 95% confidence level if $(Q) < x^2$ we accept H_0 and the selected model is an appropriate model, but if $(Q) > x^2$ we reject H_0 and we use another model because the model is not a suitable model, where both models are based on computing autocorrelation function for the residuals (Muhammad, 2011; Muhammad, 2018). However, it is probably to confirm whether the process is in the steady state statistically or not. Since the observation number is equal to one, control charts have precedence for moving range and individual values. The mean value Location CL also the upper control limit (UCL) and the lower control limit(LCL) of the ARIMA (p, d, q) control chart for individual values can be calculated from the following equations:

$$UCL = \bar{X} + \frac{3}{1.128}\bar{R}_{xi} \tag{11}$$

$$CL = \bar{X} = 0 \tag{12}$$

DOI: http://dx.doi.org/10.25098/3.1.8

Volume (3), Issue (1), Jun 2019
ISSN 2520-7377 (Online), ISSN 2520-5102 (Print)

$$LCL = \bar{X} - \frac{3}{1.128}\,\bar{R}_{xi} \qquad (13)$$

Where, $(\bar{X})$ Is the average of residual value, and $(\bar{R})$ is the average of moving range (Kovářík et al, 2015; Kovářík & Klímek, 2012; Rashid, 2016).

2.2 Tolerance limits:

Capability of ARIMA residual chart identified by the tolerance limits, to be adjusted to make the beyond points in the charts UCL and LCL, by proposing a new control limits and calculating several tolerance limits with respect to different specification values until to reach a new control limits, that makes all observations in residual charts in control. Where to compute optimal tolerance limit, specification value can be determined by taking some arbitrary constant until calculating optimal tolerance limits and selecting the nominal value close to the process mean. The tolerance limit efficiency can be tested by a measure called capability index (CP) to comparison among several specification limits and choosing a better one to be a new residual control chart for ARIMA. And capability index (CP) can be calculated as the ratio of the distance between the upper and lower specification limit (UCL-LCL) divided by (6) times the standard deviation (Ahmad, 2006).

3- Application part:

3.1 Description of the data:

The data uses in this study are taken from Kanisard factory in Sulaimani city, for producing drinking water and consists of two characteristics for the chemical water component, and these are the Power of Hydrogen (PH), Magnesium (MG), which are the most important chemical components of drinking water for (95) observations for each (PH&MG) from January to September (2018) that shows in table (2), by using the STATGRAPHICS program.

Table (2) shows the data for each (PH& MG)

n. of observation	Component of water	
1	PH	MG
2	7	7.5
3	7.1	5
4	7.3	4.7
5	7.1	8
6	7.1	8
7	7.1	5.4
8	7.2	8
9	7.1	0
10	7.2	6.5
11	7.2	6
.	.	.
.	.	.
.	.	.
91	7	3
92	6.9	5
93	7	5
94	6.8	4.6
95	7.1	7

DOI: http://dx.doi.org/10.25098/3.1.8

The Scientific Journal of Cihan University – Sulaimanyia **PP: 109-129**
Volume (3), Issue (1), Jun 2019
ISSN 2520-7377 (Online), ISSN 2520-5102 (Print)

3.2 Analysis of ARIMA control chart and tolerance chart for the chemical parameters of water (PH & MG):

3.2.1 Analysis of ARIMA residual chart and tolerance chart For (PH) process:

To determine an appropriate ARIMA model for controlling the residuals, recognizing the data behavior is an important step, for this purpose the plot of the time series is used to assess the pattern and the behaviors of data. It can be seen from the figure (1) that the (PH) process under having no trend is stationary.

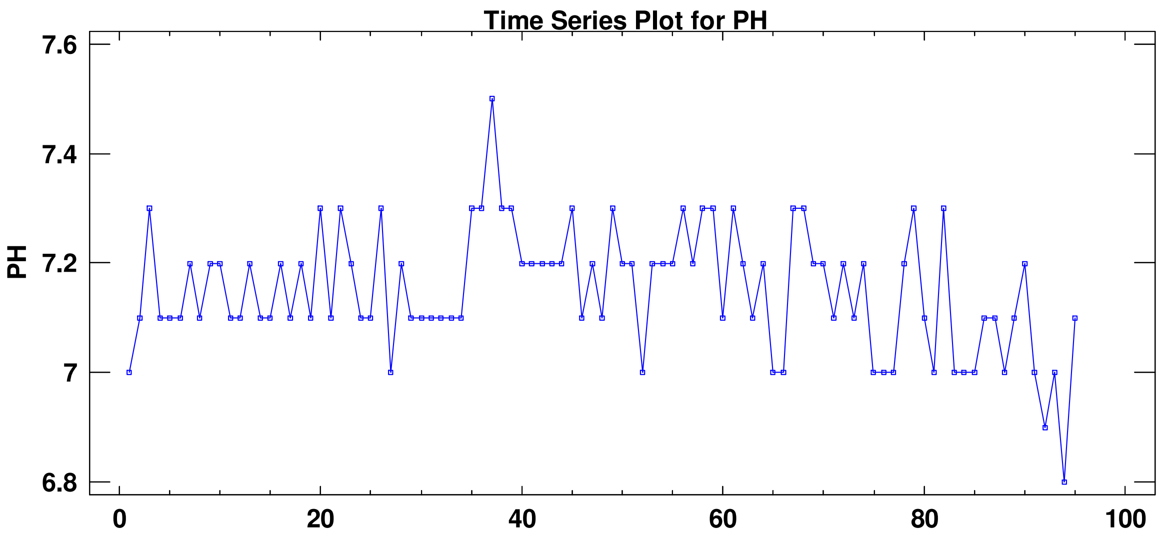

Figure (1) time series plot for (PH)

The time series plot is not enough to decide the stationary of the series, therefore for checking the stationary the autocorrelation function (ACF) and partial autocorrelation function (PACF) are plotted as shown in table (3) and (4) and figure (2) and (3), it can be seen that the value of the correlation coefficient for the series falls within the following confidence limit after lag two:

$$-1.96\frac{1}{\sqrt{n}} \leq r_k(a_t) \leq 1.96\frac{1}{\sqrt{n}} = (-0.20109, 0.20109)$$

Where (n) is the number of observation, and (k) is the lag period

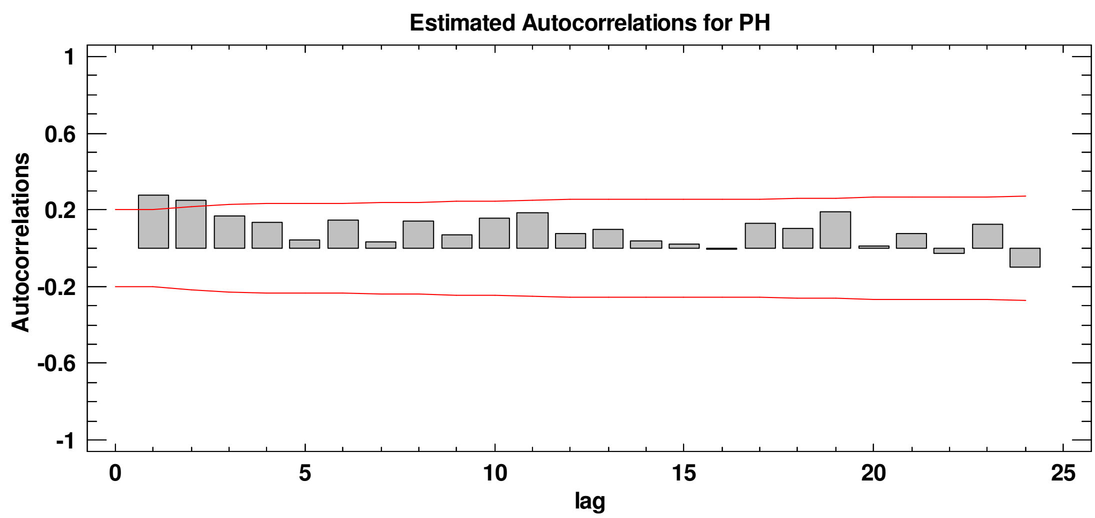

Figure (2) estimated ACF for PH

DOI: http://dx.doi.org/10.25098/3.1.8

The Scientific Journal of Cihan University – Sulaimanyia
Volume (3), Issue (1), Jun 2019
ISSN 2520-7377 (Online), ISSN 2520-5102 (Print)

Table (3) shows autocorrelation coefficient for the origin data for the (PH)

			Lower 95.0%	*Upper 95.0%*
Lag	*Autocorrelation*	*Stnd. Error*	*Prob. Limit*	*Prob. Limit*
1	0.27547	0.102598	-0.201088	0.201088
2	0.252329	0.110108	-0.215809	0.215809
3	0.170907	0.116036	-0.227426	0.227426
4	0.135284	0.118656	-0.232562	0.232562
5	0.0454234	0.120269	-0.235722	0.235722
6	0.147722	0.120449	-0.236076	0.236076
7	0.0328966	0.122341	-0.239785	0.239785
8	0.143458	0.122434	-0.239967	0.239967
9	0.0703873	0.124191	-0.24341	0.24341
10	0.155984	0.12461	-0.244232	0.244232
11	0.183036	0.126649	-0.248228	0.248228
12	0.076474	0.129403	-0.253627	0.253627
13	0.0993948	0.129878	-0.254557	0.254557
14	0.0388944	0.130677	-0.256122	0.256122
15	0.0241922	0.130798	-0.256361	0.256361
16	-0.00712395	0.130845	-0.256453	0.256453
17	0.132798	0.13085	-0.256461	0.256461
18	0.105701	0.132261	-0.259227	0.259227
19	0.19121	0.133147	-0.260964	0.260964
20	0.00957695	0.136007	-0.266569	0.266569
21	0.0783839	0.136014	-0.266582	0.266582
22	-0.0280026	0.136488	-0.267513	0.267513
23	0.124401	0.136549	-0.267631	0.267631
24	-0.0989859	0.137737	-0.269959	0.269959

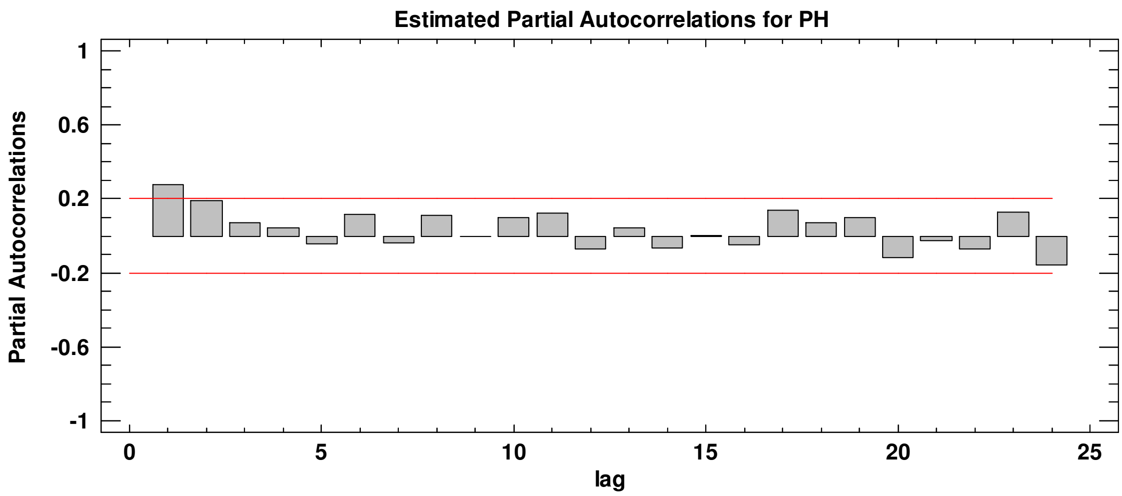

Figure (3) estimated PACF for PH

DOI: http://dx.doi.org/10.25098/3.1.8

Volume (3), Issue (1), Jun 2019
ISSN 2520-7377 (Online), ISSN 2520-5102 (Print)

Table (4) shows the partial autocorrelation coefficient for the origin data for the (PH)

Lag	*Partial Autocorrelation*	*Stnd. Error*	*Lower 95.0% Prob. Limit*	*Upper 95.0% Prob. Limit*
1	0.27547	0.102598	-0.201088	0.201088
2	0.190934	0.102598	-0.201088	0.201088
3	0.0697119	0.102598	-0.201088	0.201088
4	0.0406351	0.102598	-0.201088	0.201088
5	-0.0428851	0.102598	-0.201088	0.201088
6	0.116379	0.102598	-0.201088	0.201088
7	-0.0395987	0.102598	-0.201088	0.201088
8	0.109145	0.102598	-0.201088	0.201088
9	-0.00122465	0.102598	-0.201088	0.201088
10	0.0984715	0.102598	-0.201088	0.201088
11	0.121628	0.102598	-0.201088	0.201088
12	-0.0708304	0.102598	-0.201088	0.201088
13	0.0419056	0.102598	-0.201088	0.201088
14	-0.0658231	0.102598	-0.201088	0.201088
15	0.00445168	0.102598	-0.201088	0.201088
16	-0.0493809	0.102598	-0.201088	0.201088
17	0.136926	0.102598	-0.201088	0.201088
18	0.070426	0.102598	-0.201088	0.201088
19	0.0972204	0.102598	-0.201088	0.201088
20	-0.11768	0.102598	-0.201088	0.201088
21	-0.0247895	0.102598	-0.201088	0.201088
22	-0.0703903	0.102598	-0.201088	0.201088
23	0.128054	0.102598	-0.201088	0.201088
24	-0.158562	0.102598	-0.201088	0.201088

After checking the stationary for the series (PH) and this by studying the behavior of the autocorrelation and partial autocorrelation functions, then the next step is to determine the order of ARIMA (p, d, q) model, so by using the estimated (ACF) and estimated (PACF) plot and also after checking iteratively more than one ARIMA model with different orders as shown in table (5) the suggested models are:

DOI: http://dx.doi.org/10.25098/3.1.8

Table (5) shows the suggested models and the AKAIKE (AIC) value for (PH)

Model	σ^2	M	AIC
AR1	0.0118	1	-419.154
AR2	0.01147	2	-420437
AR3	0.01152	3	-417975
MA1	0.01216	1	-416.8654
MA2	0.01179	2	-417.830
MA3	0.01171	3	-416.4492
ARMA(1,1)	0.01125	2	-422.289
ARMA(1,2)	0.01136	3	-419.375
ARMA(1,3)	0.01142	4	-416.848
ARMA(2,1)	0.01146	3	-418.479
ARMA(2,2)	0.01149	4	-416.2534
ARMA(2,3)	0.01151	5	-414.1312

Therefore, the appropriate model is the ARMA (1, 1) model to monitor the process because this model has the minimum (AIC) among all the suggested models which are equal to (-422.289). The estimated model for the series (PH) for ARMA (1, 1) is:

$$X_t = 0.231913 + 0.967438x_{t-1} + \varepsilon_t - 0.817773\varepsilon_{t-1} \tag{14}$$

Where X_t is predicted value, all values are coefficients, and ε_t is residual for period (t).

After selecting the appropriate model, then the final step is checking the estimated residuals for the ARMA (1, 1) model, for this purpose the (ACF) and (PACF) plots of the estimated residuals for (PH) series are constructed respectively in figure (4) and (5) to determine whether there is any autocorrelation occurs in the residuals for ARMA (1, 1) model.

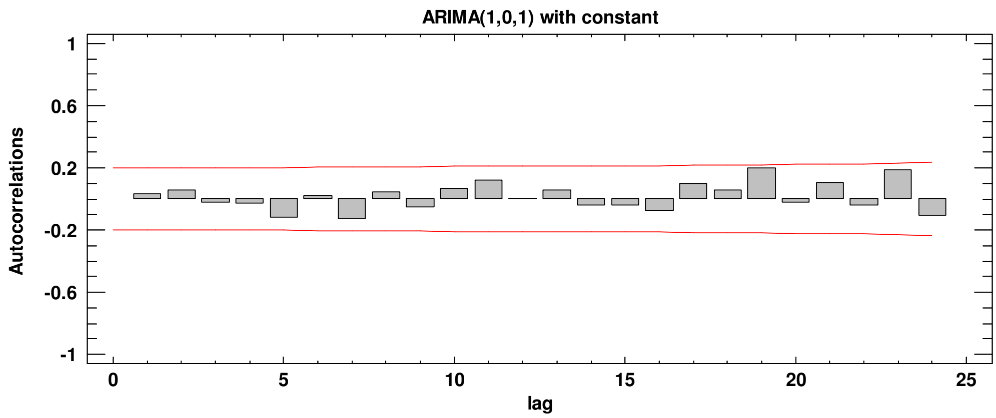

Figure (4) ACF plot for estimated residual

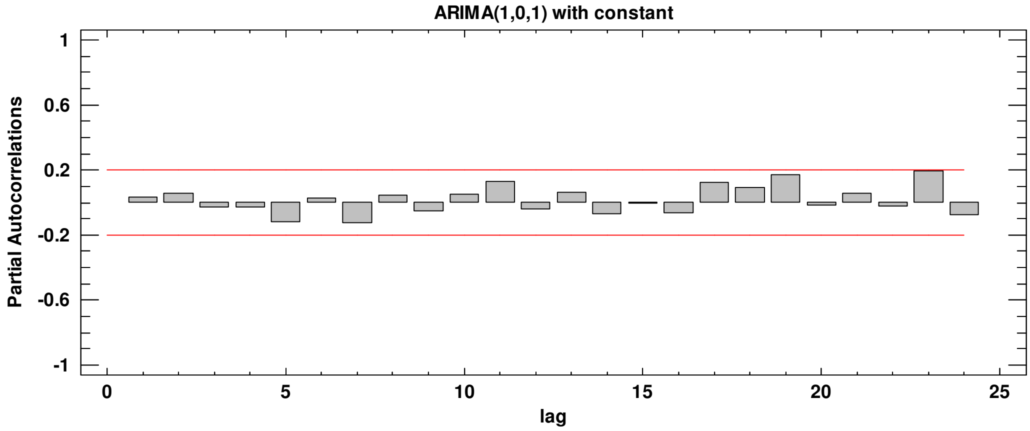

Figure (5) PACF plot for estimated residual

From plotting (ACF) and (PACF) for selected ARMA (1, 1) model for the (PH) data, it was shown that all the coefficients fall within the following confidence limits:

$$-1.96\frac{1}{\sqrt{n}} \leq r_k(a_t) \leq 1.96\frac{1}{\sqrt{n}} = (-0.20109, 0.20109)$$

And this shows that the suggested model is an appropriate model, and this result can be supported by the Box-Pierce test comparing with the chi-square (x^2) table.

$H_0: p_k(a) = 0$ V/s $H_1: p_k(a) \neq 0$

$$Q = n\sum_{k=1}^{n} r^2(a\hat{}) = 17.422 < x^2(22) = 36.42 \quad (15)$$

By comparing (Q) statistics with the $chi - square(x^2)$ table and (%95) confidence level and (22) degree of freedom, the (Box-Pierce) test result revealed that the estimated residuals are independent, therefore the ARMA (1, 1) model is the suitable model for (PH) process data.

Table (6) gives the summary for the estimated parameters of the ARIMA (1,0,1) model, and ARIMA residual chart with the number of the out of control point, also shows the process capacity for the (PH) process:

Table (6) shows summary for ARIMA chart & process capacity for (PH)

A	ARIMA model summary Parameter estimate Constant 0.231913 AR(1) 0.967438 MA(1) 0.817773	C	Estimates Process mean= 7.12213 Process sigma=0.117416
B	ARIMA chart UCL: +3.0 sigma= 0.303206 Center line=0.0 LCL: -3.0 sigma= -0.303206 1 beyond limits	D	Capability indices for PH Specifications USL=7.47 Nominal=7.12 LSL=6.77 Cp=0.9936 Cpk=0.9875 Cpk(upper)=0.9875 Cpk(lower)=0.9996

DOI: http://dx.doi.org/10.25098/3.1.8

Volume (3), Issue (1), Jun 2019
ISSN 2520-7377 (Online), ISSN 2520-5102 (Print)

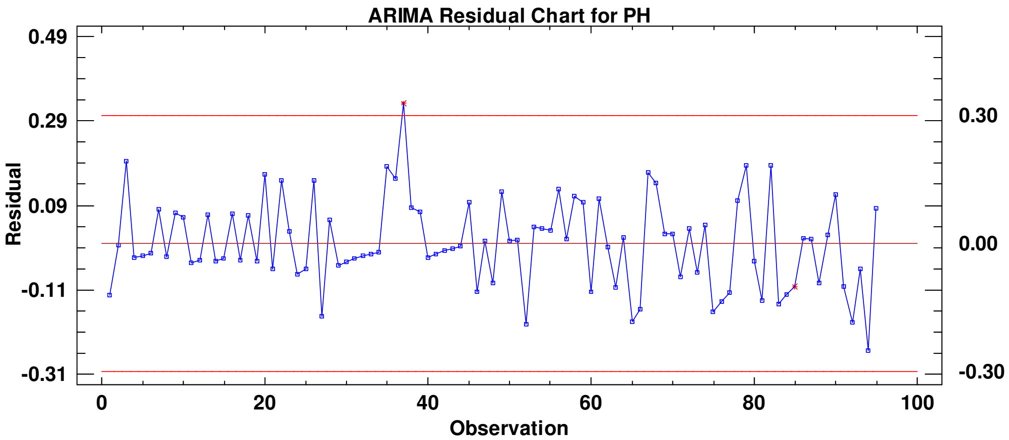

Figure (6) ARIMA control chart for (PH)

Under the condition of normality and randomness of the residual series generated from ARIMA (1,0,1) which shows only one point is beyond the limit and this is random and does not have the negative effect of the (PH) process, therefore the (PH) process is in statistical control.

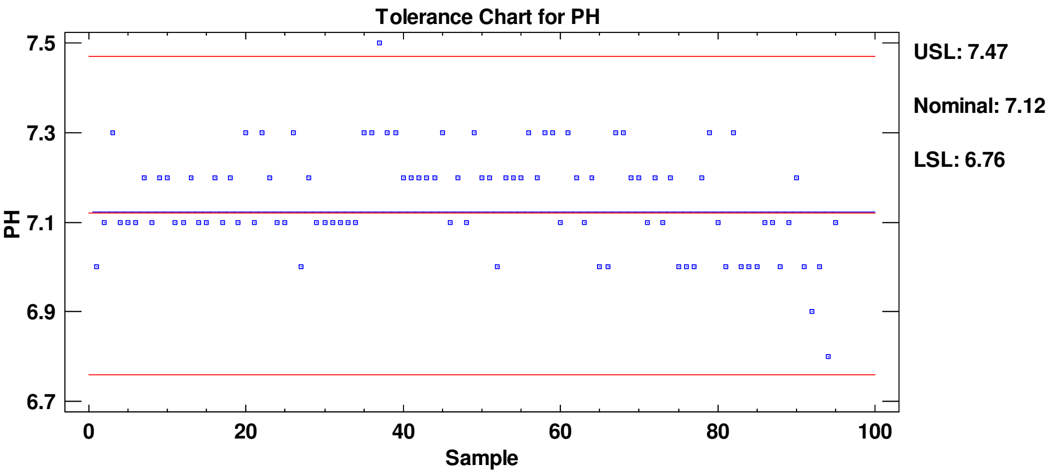

Figure (7) Tolerance chart for (PH)

From the table (6) it is shown that the (PH) process is not capable and it can be seen by the capability index (Cp) which is equal to (0.9936) for the control limit (6.77, 7.47), therefore from comparing several tolerance limit for (PH) the optimal tolerance limit for (PH) with remaining the process is in control is equal to (6.76, 7.47) which has the capability index is equal to (1.00782) and this shows that the suggested control limit for (PH) process is a good control limit. Therefore, the factory should make an adjustment from the control limit (6.77, 7.47) to (6.76, 7.47) for producing the water in the better way, also tolerance limit that used for constructing a new structure control chart for controlling (PH) process does not allow more than (7.47).

3.2.2 ARIMA residual chart and tolerance chart For (MG) process:

To determine an appropriate ARIMA model for controlling the residuals, recognizing the data behavior is an important step, for this purpose the plot of the time series is used to assess the pattern and the behaviors of data. It can be seen from the figure (8) that the (MG) process under having no trend is stationary.

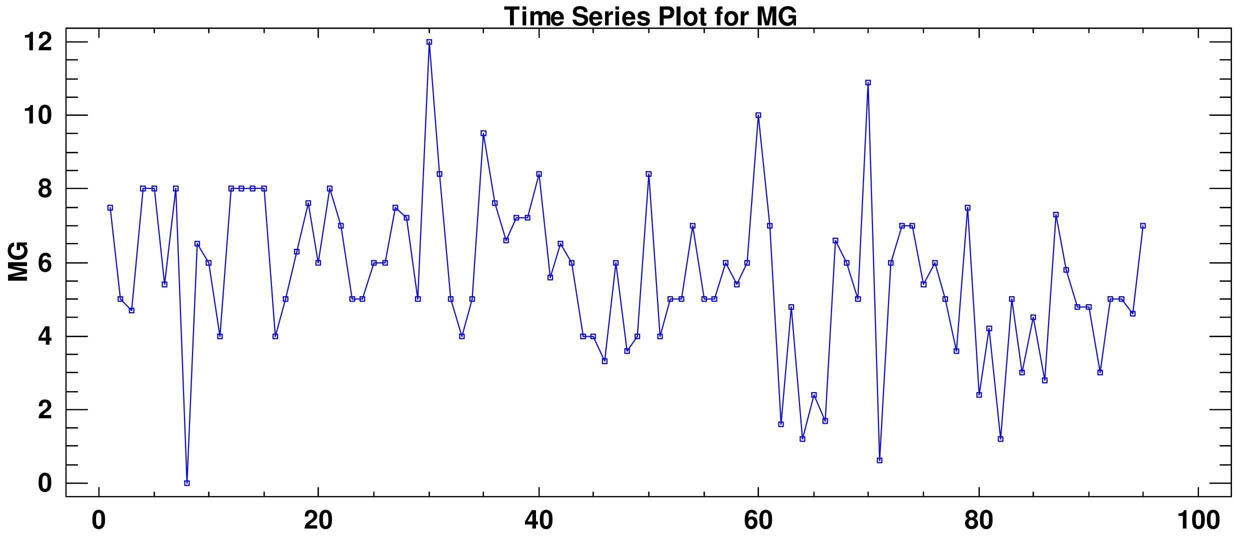

Figure (8) time series plot for (MG)

The time series plot is not enough to deciding the stationary of the series, therefore for checking the stationary the autocorrelation function (ACF) and partial autocorrelation function (PACF) are plotted as shown in table (7) and (8) and figure (9) and (10), it can be seen that the value of the correlation coefficient for the series falls within the following confidence limit:

$$-1.96\frac{1}{\sqrt{n}} \leq r_k(a_t) \leq 1.96\frac{1}{\sqrt{n}} = (-0.20109, 0.20109)$$

Where (n) is the number of observation, and (k) is the lag period.

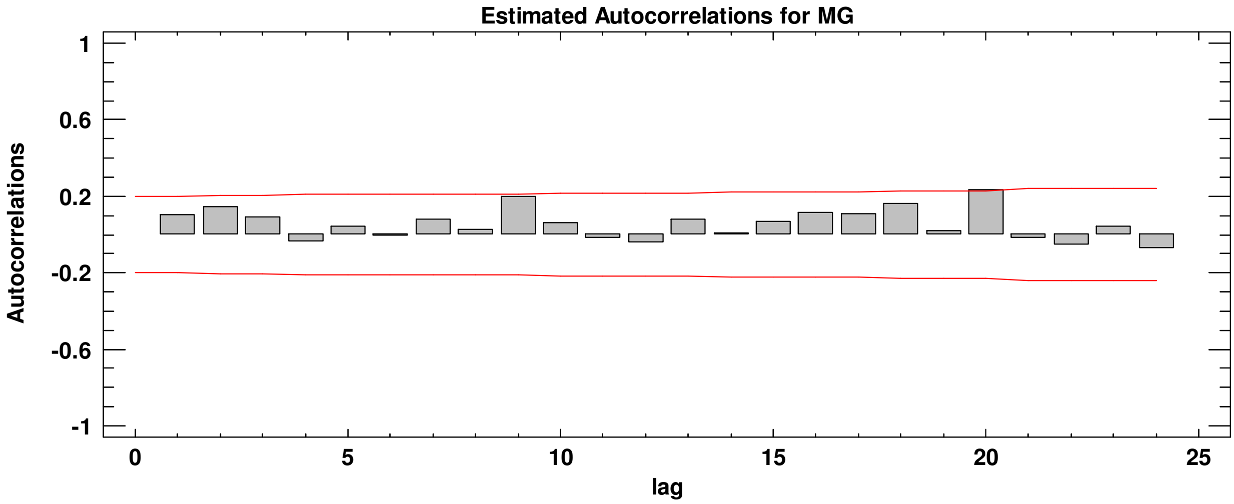

Figure (9) Estimated ACF for MG

DOI: http://dx.doi.org/10.25098/3.1.8

Volume (3), Issue (1), Jun 2019
ISSN 2520-7377 (Online), ISSN 2520-5102 (Print)

Table (7) shows autocorrelation coefficient for the origin data for the (MG)

Lag	Autocorrelation	Stnd. Error	Lower 95.0% Prob. Limit	Upper 95.0% Prob. Limit
1	0.102006	0.102598	-0.201088	0.201088
2	0.144519	0.10366	-0.20317	0.20317
3	0.0915181	0.105759	-0.207285	0.207285
4	-0.0308189	0.10659	-0.208913	0.208913
5	0.0421821	0.106684	-0.209096	0.209096
6	-0.000389077	0.106859	-0.20944	0.20944
7	0.0798009	0.106859	-0.20944	0.20944
8	0.029602	0.107485	-0.210666	0.210666
9	0.200929	0.10757	-0.210834	0.210834
10	0.0616173	0.111451	-0.21844	0.21844
11	-0.01664	0.111809	-0.219142	0.219142
12	-0.0411777	0.111835	-0.219193	0.219193
13	0.0781915	0.111995	-0.219506	0.219506
14	0.00904605	0.112568	-0.220629	0.220629
15	0.0687976	0.112575	-0.220644	0.220644
16	0.11848	0.113017	-0.22151	0.22151
17	0.108522	0.114317	-0.224058	0.224058
18	0.160884	0.115396	-0.226173	0.226173
19	0.0187114	0.117734	-0.230754	0.230754
20	0.234222	0.117765	-0.230816	0.230816
21	-0.0125628	0.122571	-0.240234	0.240234
22	-0.0503346	0.122584	-0.240261	0.240261
23	0.046996	0.122802	-0.240687	0.240687
24	-0.0654075	0.122991	-0.241058	0.241058

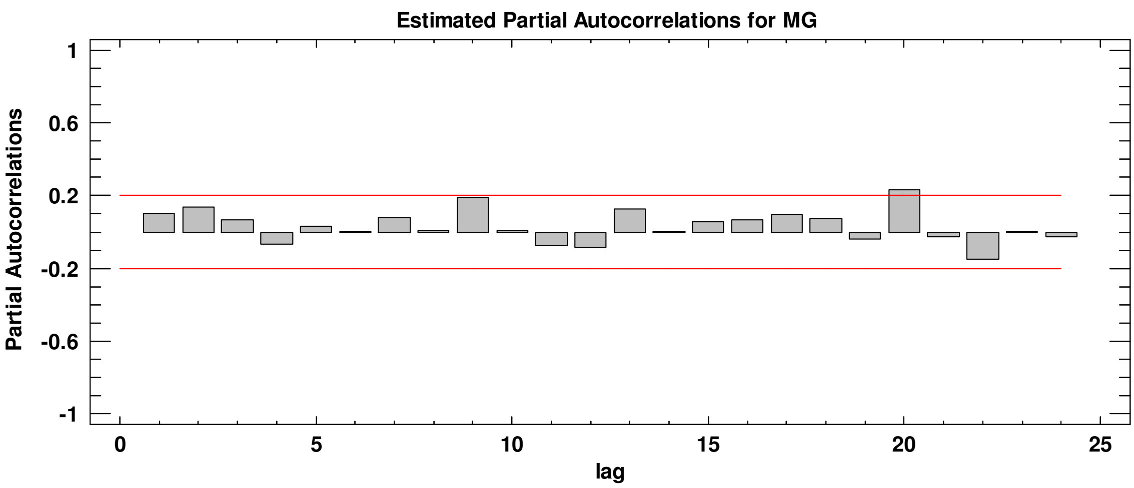

Figure (10) Estimated PACF for MG

Table (8) shows the partial autocorrelation coefficient for the origin data for the (MG)

Lag	Partial Autocorrelation	Stnd. Error	Lower 95.0% Prob. Limit	Upper 95.0% Prob. Limit
1	0.102006	0.102598	-0.201088	0.201088
2	0.135524	0.102598	-0.201088	0.201088
3	0.066861	0.102598	-0.201088	0.201088
4	-0.0657793	0.102598	-0.201088	0.201088
5	0.0295023	0.102598	-0.201088	0.201088
6	0.00086174	0.102598	-0.201088	0.201088
7	0.0811833	0.102598	-0.201088	0.201088
8	0.00876838	0.102598	-0.201088	0.201088
9	0.187893	0.102598	-0.201088	0.201088
10	0.011135	0.102598	-0.201088	0.201088
11	-0.0720574	0.102598	-0.201088	0.201088
12	-0.0851833	0.102598	-0.201088	0.201088
13	0.126622	0.102598	-0.201088	0.201088
14	0.00552175	0.102598	-0.201088	0.201088
15	0.0538016	0.102598	-0.201088	0.201088
16	0.0666046	0.102598	-0.201088	0.201088
17	0.0943051	0.102598	-0.201088	0.201088
18	0.0754683	0.102598	-0.201088	0.201088
19	-0.0405466	0.102598	-0.201088	0.201088
20	0.232106	0.102598	-0.201088	0.201088
21	-0.027971	0.102598	-0.201088	0.201088
22	-0.151114	0.102598	-0.201088	0.201088
23	0.00334161	0.102598	-0.201088	0.201088
24	-0.0278361	0.102598	-0.201088	0.201088

After checking the stationary for the series (MG) and this by studying the behavior of the autocorrelation and partial autocorrelation functions, then the next step is to determine the order of ARIMA (p, d, q) model, so by using the estimated (ACF) and estimated (PACF) plot and also after checking iteratively more than one ARIMA model with different orders as shown in table (9) the suggested models are:

Therefore, the appropriate model is the ARIMA (1,0,0) model to monitor the process because this model has the minimum (AIC) among all the suggested models which are equal to (146.92). The estimated model for the series (MG) for ARIMA (1,0,0) is:

$$X_t = 5.07424 + 0.102893x_{t-1} + \varepsilon_t \tag{16}$$

Where X_t is predicted value, all values are coefficients, and ε_t is residual for period (t).

DOI: http://dx.doi.org/10.25098/3.1.8

Table (9) shows the suggested models and the AKAIKE (AIC) value for (MG)

Model	σ^2	M	AIC
AR1	4.5976	1	146.92
AR2	4.5617	2	148.18
AR3	4.5915	3	150.799
MA1	4.6081	1	147.142
MA2	4.5680	2	148.312
MA3	4.58021	3	150.56
ARMA(1,1)	4.64812	2	149.963
ARMA(1,2)	4.60679	3	151.115
ARMA(1,3)	4.6325	4	153.644
ARMA(2,1)	4.61383	3	151.260
ARMA(2,2)	4.65892	4	154.184

After selecting the appropriate model, then the final step is checking the estimated residuals for the ARIMA (1,0,0) model, for this purpose the (ACF) and (PACF) plots of the estimated residuals for (MG) series are constructed respectively in figure (11) and (12) to determine whether there is any autocorrelation occurs in the residuals for ARIMA (1,0,0) model.

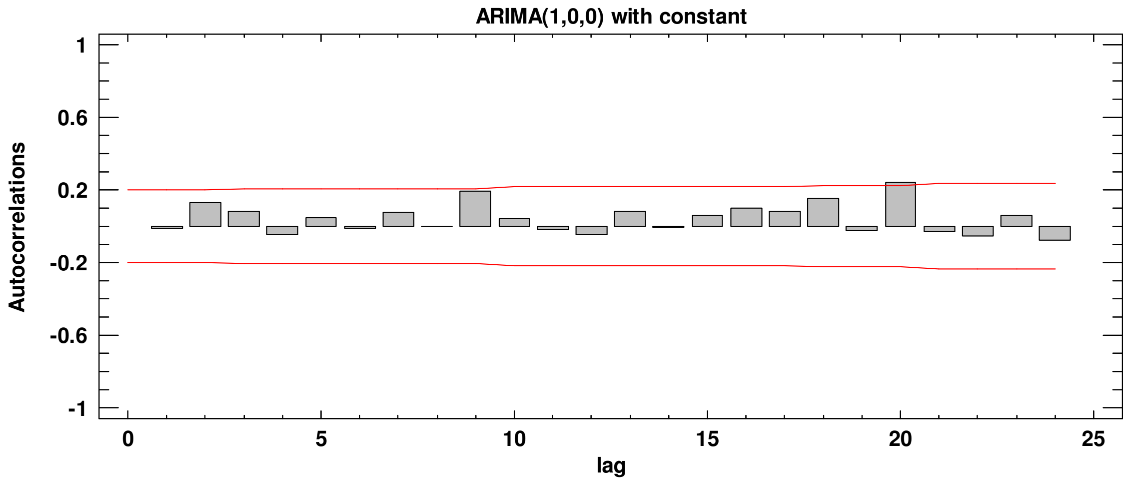

Figure (11) ACF plot for estimated residual

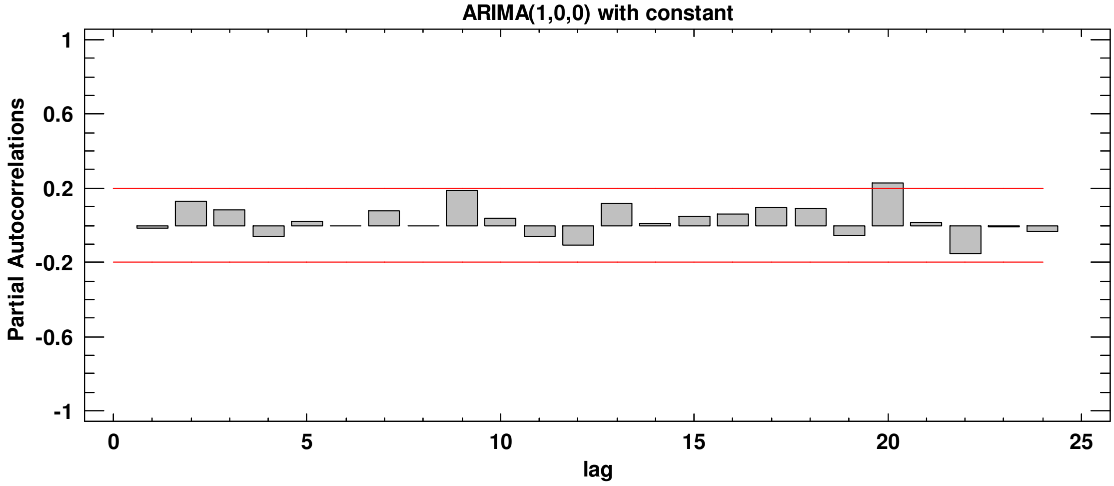

Figure (12) PACF plot for estimated residual

From plotting (ACF) and (PACF) for selected ARIMA (1,0,0) model for the (MG) data, it is shown that all the coefficients fall within the following confidence limits:

$$-1.96\frac{1}{\sqrt{n}} \leq r_k(a_t) \leq 1.96\frac{1}{\sqrt{n}} = (-0.20109, 0.20109)$$

And this shows that the suggested model is an appropriate model, and this result can be supported by the Box-Pierce test comparing with the chi-square (x^2) table.

$H_0: p_k(a) = 0$ V/s $H_1: p_k(a) \neq 0$

$$Q = n \sum_{k=1}^{n} r^2(\hat{a}) = 18.7878 < x^2(23) = 36.42 \tag{17}$$

By comparing (Q) statistics with the $chi - square$ (x^2) table and (%95) confidence level and (23) degree of freedom, the (Box-Pierce) test result revealed that the estimated residuals are independent, therefore the ARIMA (1,0,0) model is the suitable model for (MG) process data.

Table (10) gives the summary for the estimated parameters of the ARIMA (1,0,0) model, and ARIMA residual chart with the number of the out of control point, also shows the process capacity for the (MG) process:

Table (10) shows summary for ARIMA chart & process capacity for (MG)

A	ARIMA model summary Parameter estimate Constant 5.07424 AR(1) 0.102893	C	Estimates Process mean= 5.65623 Process sigma=1.99443
B	ARIMA chart UCL: +3.0 sigma= 5.95152 Center line=0.0 LCL: -3.0 sigma= -5.95152 1 beyond limits	D	Capability indices for MG Specifications USL=11.64 Nominal=5.66 LSL=-0.33 Cp=1.00029 Cpk=1.00008 Cpk(upper)=1.00008 Cpk(lower)=1.00049

DOI: http://dx.doi.org/10.25098/3.1.8

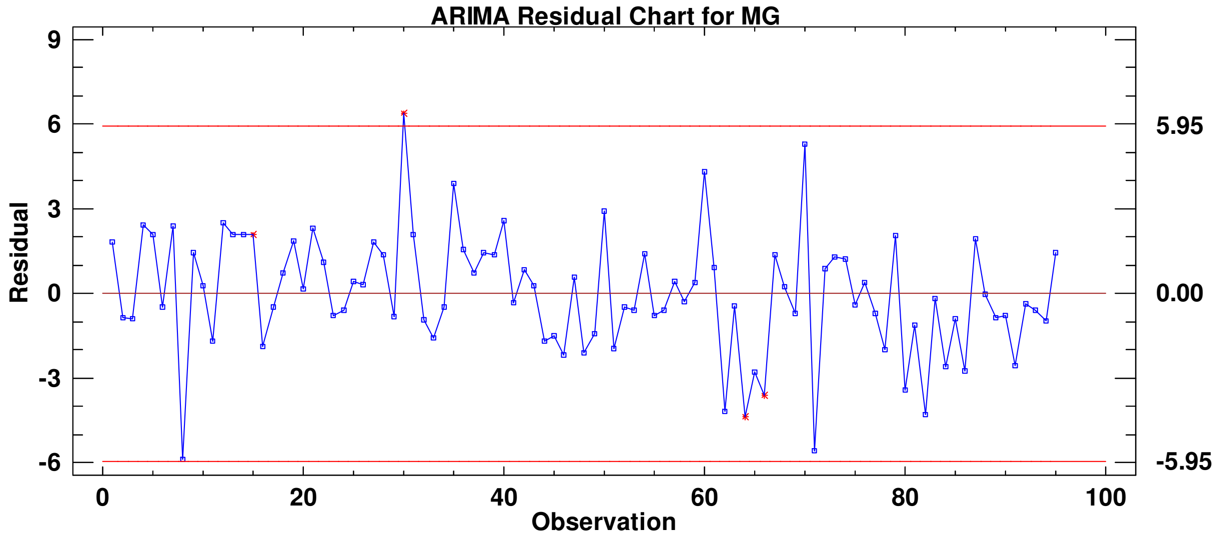

Figure (13) ARIMA control chart for (MG)

Under the condition of normality and randomness of the residual series generated from ARIMA (1,0,0) which shows one point is beyond the limit and this is random and does not have the negative effect of the (MG) process, therefore the (MG) process is in statistical control.

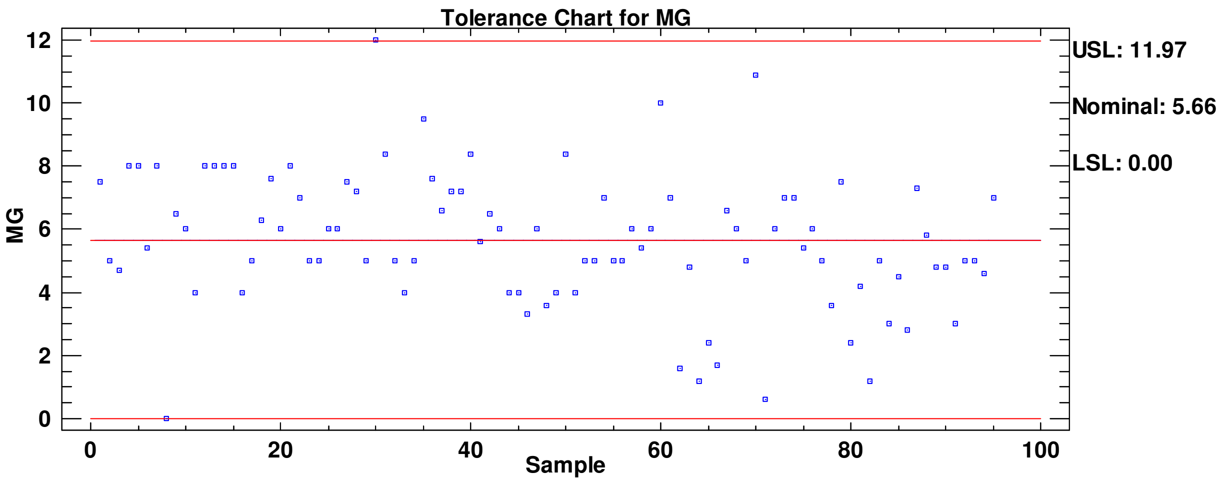

Figure (14) Tolerance chart for (MG)

From table (10) it is shown that the (MG) process is capable and the (Cp) is equal to (1.00029) but not has the optimal tolerance limits which is equal to (-0.33, 11.64), therefore from comparing several tolerance limits for (MG) process, the optimal tolerance limit for (MG) with remaining the process is in control is equal to (0, 11.97) which has the capability index is equal to (1.00029) and this shows that the suggested control limit for (MG) process is a good control limit. Therefore, the factory should make an adjustment from the control limit (-0.33, 11.64) to (0, 11.97) for producing the water in a better way.

4- Conclusion:

Controlling and monitoring the production process is very important for improving the quality of the product and reducing the costs, also using the statistics in quality control is a very important issue for improving the product quality or keeping the process in the acceptance region by using the most important tool the control charts. This study aimed to monitor the two important chemical parameters of drinking water and these are Power of Hydrogen (PH) and Magnesium (MG), to determine if they are out of control or in control and also determining the control limits for both (PH &MG) from the optimal tolerance limits to control the water production to reach better quality products in the future, by taking the data for each of the parameters (PH& MG) from the (KANISARD) factory of drinking water at (Sulaimani) city in Kurdistan region in Iraq. To achieve this purpose, ARIMA control charts were constructed for each of the parameters, where ARIMA residual chart is a special control chart, used to specify and detect the quality behavior in time-correlated process data, and this type of chart is also useful for adjusting and specifying the quality limits during the process. The result of the study showed that the estimated ARMA (1, 1) and ARIMA (1, 0, 0) model is an adequate model for both (PH) and (MG) behaviors respectively for constructing an identically and normally distributed residuals series. And also the models were very helpful to achieve a non-auto-correlated residual series, which is an important tool to fit a suitable ARIMA residual control charts for (PH) and (MG). By using the selected models of the ARIMA residual charts for both (PH &MG) respectively it was concluded that both (PH& MG) process is in statistical control, in addition, from the tolerance limits and capability index, the suggested control limit for both (PH) and (MG) respectively are (UCL, LCL for PH) = (6.76, 7.47), and (UCL, LCL for MG) = (0, 11.97) for controlling the water production and to reach the better quality products in the future.

References:

Ashour, M. F. M. (2014). Using multivariate statistical quality control models to monitor the quality of drinking water in KHAN YOUNIS governorate. M.S.C thesis, department of applied statistics, faculty of economics and administrative sciences, AL-AZHER University

Ahmad, N. M. (2006). Constructing control charts, for cigarettes production, in Sulaimani cigarettes factory. Journal of administration and economics

Bhasin, S., Shukla, A. N., Shrivastava, S., & Mishra, U. (2016). Control Chart Model for Assessment of Water Quality of a Tropical River-Kshipra Ujjain, India.

George, J. P., Chen, Z., & Shaw, P. (2009). Fault detection of drinking water treatment process using PCA and hotelling's T^2 chart. *World Academy of Science, Engineering and Technology*, *50*, 970-975.

Hamarasoul, F. O., & Hidar, E. A. (2019). Use of the modified and standard deviation plates to control the quality property of amoxicillin and predict the control residue of the time series model. *Journal of Humanity Sciences*, *23*(1), 1-24.

Kovářík, M., Sarga, L., & Klímek, P. (2015). Usage of control charts for time series analysis in financial management. *Journal of Business Economics and Management*, *16*(1), 138-158.

Kovářík, M., & Klímek, P. (2012). The usage of time series control charts for financial process analysis. Journal of Competitiveness

 DOI: http://dx.doi.org/10.25098/3.1.8

Magaji, A. A., Yahaya, A., & Asiribo, O. E. (2015). Assessing the Effects of Autocorrelation on the Performance of Statistical Process Control Charts. International Journal of mathematics and statistics invention, 3(6), 15-23.

Montgomery, D. C. (2007). Introduction to statistical quality control. John Wiley & Sons

Muhammad, M. A. (2011). The analysis and forecasting in time series

Muhammad, M. A. (2018). Time series analysis with application

Rashid, K. M. J. (2016). The Effect of Sample Size On (Cusum and ARIMA) Control Charts. International Journal of Advanced Engineering, Management and Science, 2(5)

Salleh, R. M., Zawawi, N. I., Gan, Z. F., & Nor, M. E. (2018). Autocorrelated process control: Geometric Brownian Motion approach versus Box-Jenkins approach. In *Journal of* Physics: Conference Series (Vol. 995, No. 1, p. 012039). IOP Publishing

Salih, S. O. (2011). Using Fuzzy Logic in Quality Control on ALA natural water production in Sulaimani. M.S.C Thesis College of administration and economics, University of Sulaimani

TAŞDEMİR, A. (2017). DETERMINATION OF CONTROL LIMITS FOR ASH CONTENT OF CLEAN COARSE COAL PROCESSED BY HEAVY MEDIUM DRUM. The Online Journal of Science and Technology-April, 7(2)

Tasdemir, A., & Kowalczuk, P. B. (2014). Application of statistical process control for proper processing of the Fore-Sudetic Monocline copper ore. Physicochemical Problems of Mineral Processing, 50

Xiao, P. (2013). Robust MEWMA-type control charts for monitoring the covariance matrix of multivariate processes (Doctoral dissertation, Virginia Tech).