IRAQI Academic Scientific Journals

University of Kirkuk Journal For Administrative and Economic Science (2025) 15 (3) Part (1): 281-288

UKJAES

University of Kirkuk Journal For Administrative and Economic Science

ISSN:2222-2995 E-ISSN:3079-3521
University of Kirkuk Journal For Administrative and Economic Science

Hasan Samyia Khalid & Ahmad Goran Zyad. Using Regression Control Charts to Forecast Garota Factory in Erbil, Kurdistan Region. *University of Kirkuk Journal For Administrative and Economic Science* **(2025) 15 (3) Part (1):281-288.**

Using Regression Control Charts to Forecast Garota Factory in Erbil, Kurdistan Region

Samyia Khalid Hasan [1], Goran Zyad Ahmad [2]

[1,2] *College of Administration & Economics, Salahaddin University, Erbil, Iraq*

saiya.hasan@su.edu.krd [1]
goranstat@gmail.com [1]

Abstract: Wheat is one of the most widely grown cereal grains in the world. It is a staple food for the world’s population. A quality for regression model is an effective tool to forecast statistical quality control used to monitor and assess the performance and stability of the model over time. In this study investigates the application of statistical process control in order to maintain produced of quality. The data set of this study was obtained from Garota flour factory in Erbil between January in 2019 to December 2024 by months, our goal is using regression control charts for the number of working hours at Garota factory to forecast flour outcomes by quantity, detecting deviations, outliers, and then addressing them to achieve optimal performance. The result shows that regression control chart is powerful of using to quality protected. Here it can see that the whole process is under control. the results obtained using statistical packages IBM (SPSS-28 and STATGRAPHICS-19).

Keywords: Statistical Process Control (SPC), Regression Analysis, Residual Control Charts, Average Chart (X-Bar Chart), Standard Deviation (S.D Chart).

استخدام مخططات التحكم في الانحدار للتنبؤ بمصنع غاروتا في أربيل، إقليم كوردستان

أ.م.د. سامية خالد حسن [١]، الباحث: گۆران زياد احمد [٢]

[١،٢] *كلية الادارة والاقتصاد، جامعة صلاح الدين، أربيل، العراق*

المستخلص: القمح هو أحد أكثر الحبوب انتشارًا في العالم. إنه غذاء أساسي لسكان العالم.يُعد نموذج الانحدار أداةً فعّالة للتنبؤ بمراقبة الجودة الإحصائية، ويُستخدم للمراقبة وتقييم أداء النموذج واستقراره مع مرور الوقت. تطبيق على هذه الدراسة مراقبة العمليات الإحصائية للحفاظ على جودة الإنتاج. تم جمع البيانات لدراسة في مصنع دقيق غاروتا في أربيل خلال فترة الزمنية من يناير ٢٠١٩ إلى ديسمبر ٢٠٢٤. هدفنا هو استخدام مخططات مراقبة الانحدار لعدد ساعات العمل في مصنع غاروتا وتنبؤ بنتائج الدقيق حسب الكمية، واكتشاف الانحرافات والقيم المتطرفة، ثم معالجتها لتحقيق الأداء الأمثل. تبين من خلال النتائج بانة عند استخدام لوحة الانحدار للسيطرة كانت افضل لحماية الجودة، تظهر أن العمليات الانتاجية تحت السيطرة. تم الحصول على النتائج باستخدام حزمتي الإحصائيتين (SPSS-28 و STATGRAPHICS-19).

الكلمات المفتاحية: التحكم الإحصائي في العملية (**SPC**)، تحليل الانحدار، البواقي المخططات التحكم، لوحة المتوسط (لوحة **X-Bar**)، لوحةالانحراف المعياري). لوحة **S.D**).

Corresponding Author: E-mail: saiya.hasan@su.edu.krd

Introduction

A Regression Control Chart is a statistical quality control tool used to monitor processes where the quality characteristic of interest is influenced by one or more independent (predictor) variables. (Hasan, 2018). Unlike traditional control charts-which assume the process mean is constant over time-regression control charts adjust for systematic changes due to these influencing variables, using regression models to account for and remove variation caused by external factors. This makes them especially valuable in industrial, engineering, and research settings where process behavior is influenced by controllable variables like temperature, speed, or pressure. Their main objective was to improve process monitoring by removing predictable, non-random variation using linear regression, thus isolating the true random variation that control charts are designed to detect. (Al-Mashhadani,2015; Montgomery, 2009).

DiPaola was the first to propose the concept of a regression control chart (1945). which is a statistical technique used to analyze relationships between two variables. These charts are a type of control chart that uses a regression line (a line representing the relationship between two variables) as the central tendency, and control limits (upper and lower bounds) are placed around the regression line, The goal of a regression control chart is to monitor the process stability, ensuring that the relationship between the two variables stays within the defined limits. This method was similar to that described by DiPaola (1945) and aimed to determine if a process was stable or out of control based on the relationship between the two variables, The use of tolerance limits as control limits is clearly preferred by Wallis and Roberts (1956) and Mandel (1969), who offer a chart with a regression line and confidence intervals as control limits. The regression line serves to determine the mean value, and it is assumed that the y values are independently and normally distributed. The variations between the actual and expected y values from the regression line are used to determine the standard error, se, which is independent of the values of x. When k is a constant (usually k = 2 or 3), the Regression Control Chart appears like this: $\hat{y} \pm kse$. Control limits set to k=2 it means that the reveal a 95.45 % confidence interval and k=3 reveal a 99.7% confidence interval. (Ghosh, ather1981; Ghosh and Reynolds,1981; Besterfield, 1979,2000).

1st: Quality Control Charts

A quality control chart plots data points against set control limits to help learn variances and maintain consistency in production. It is a statistical tool used to track and control a process over time. Bell Telephone Laboratories' Walter A. Shewhart devised the first quality control chart in 1924, and he and his coworker continued to refine it. In 1931, he gave a thorough explanation of control charts. Shewhart's concept: Shewhart built his charts using the normal distribution, and the statistical grounds for the formation and usage of quality assurance graphs were supplied by statistical inference and sampling theory. He found that a distribution may be converted into a normal shape by estimating its mean and standard deviation. Shewhart had to create his charts using the normal distribution since he wanted to know if the production process was moving smoothly and naturally and if the points on the chart matched a normal distribution. A stable distribution is defined as one in which, for the process under statistical control, variation does not go beyond the predetermined bounds more than 0.27 percent of the time. Shewhart applied the central limit theorem, which includes three components related to control charts, to define a stable distribution: The distribution of averages from a population's individual item averages produces a normal distribution curve, which is also known as a bell shape. (Kadir, et al, 2024) seen in Fig.1(a).

three parallel lines make up Shewhart control charts and they are as follows:

1. The mean or overall average, of the quality feature being monitored, is represented by the letter T as the center line (also known as the target line) of the control chart.

the center limits control (CCL) like with Average Chart (X- Bar Chart) chart is used to monitor and analyze changes in the process average (mean) over time. This chart's goal line is the overall

average of all sample means, which organizations may use to ensure that their processes are reliable, consistent, and within acceptable parameters and is calculated using the following formula:
Where: $\bar{\bar{x}}$: average of the sample mean, $\bar{x}_j$: average of the subgroup, m: number of samples.
2. For a process in a state of control the upper control limit is the greatest allowable deviation from the mean. represented mathematically as $UCL = \bar{\bar{x}} + 3\sigma$
3. For a process in a state of control the lower control limit is the greatest allowable deviation from the mean. represented mathematically as $LCL = \bar{\bar{x}} - 3\sigma$
where the vertical axis shows the quality include and the horizontal axis shows the order of samples (or time). The making decision process in qualitative control Charts is a speedy and alternative means of frequently.

testing the following hypothesis:

$H0 : \mu = T$, $H1 : \mu \neq$ T. Where (μ): The average is the statistical pointer of the qualitative characteristic to be controlled. $H0$: This means that the production process is under statistical control. It can be shown in Fig.1(b). (Ghosh and Yer, 1981)

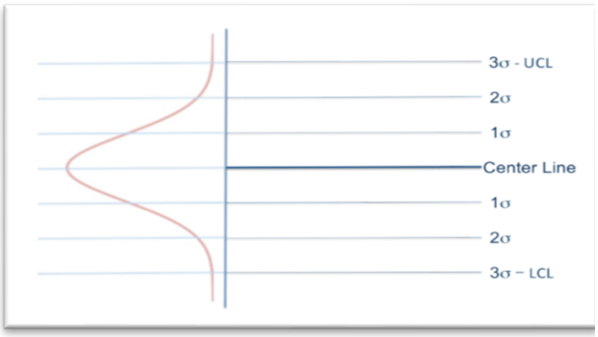

Fig.1(a): Normal distribution

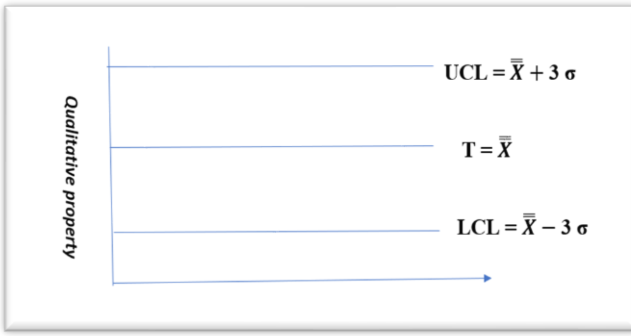

Fig.1(b): Average Chart Sample sequence

2nd: Mathematics of the Regression Control Chart

The Regression Control Chart is based on using the center line (CL) with upper control limits (UCL) and lower control limits (LCL) to monitor and regulate an independent variable from simple linear regression. The goal of simple linear regression is to use a collection of data to determine the line that best fits them. Using the least squares approach, the regression equation's coefficients are found. It minimizes the sum of the squared residuals from the regression line. (Al-Taie and Al-Hamdani, 2018; Aslan and Gunhan 2014).

Building a fundamental linear regression model is to build a Regression Control Chart. This model is made as follows: (Baris and Aslan, 2011;).

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \quad , i = 1,2,3,\dots, n \qquad (1)$$

where: N: sample size, Y_i : dependent variable, X_i : independent variable, β_0 : intercept point, β_1 : slope regression line, ε_i : random error. (Montgomery, 2005)
Building a fundamental linear regression model for the information is the first step in building a control regression panel. This model is made as follows:
where: N: sample size, Y_i : dependent variable, X_i : independent variable, β_0 : intercept point, β_1 : slope regression line, ε_i : random error. (Montgomery, 2005)
Building a fundamental linear regression model for the information is the first step in building a control regression panel. This model is made as follows:

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i \qquad (2)$$

The error value is the difference between the actual and estimated value of the dependent variable, is given according to the following formula: (Kadir and Jamil,2024).

$e_i = Y_i - \hat{Y}_i$, i = 1,2,3,…, n (3)

Sum squared residuals $\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n}[Y_i - \hat{Y}_i]^2 = \sum_{i=1}^{n}[Y_i - (\hat{\beta}_0 + \hat{\beta}_1 X_i)]^2$

$$E(\beta_0, \beta_1) = \sum_{i=1}^{n}[Y_i - (\hat{\beta}_0 + \hat{\beta}_1 X_i)]^2$$

The partial derivatives equal to zero and find solutions:

$$\frac{\partial E(\beta_0, \beta_1)}{\partial \beta_0} = 0 \ , \ \frac{\partial E(\beta_0, \beta_1)}{\partial \beta_1} = 0$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} \quad (4)$$

The mean and standard deviation can be found according to the following two methods:

$$\bar{e} = \frac{\sum_{i=1}^{n} e_i}{n} \quad (5)$$

$$\sigma_e = \sqrt{MS_{Residual\ Error}} \ , \ \sigma_e = \sqrt{\frac{\sum_{i=1}^{n}(e_i - \bar{e})^2}{n-1}} \quad (6)$$

The last step is to find the limits of the Regression Control Chart, which is given by the following formula:

$$UCL = \hat{\beta}_0 + \hat{\beta}_1 X_i + k\ \sigma_e$$
$$CCL = \hat{\beta}_0 + \hat{\beta}_1 X_i \quad (7)$$
$$LCL = \hat{\beta}_0 + \hat{\beta}_1 X_i - k\ \sigma_e$$

Where k is a constant (typically k=2 or 3). Control limits set to k=3 equivalent to a % 99.7 confidence interval.

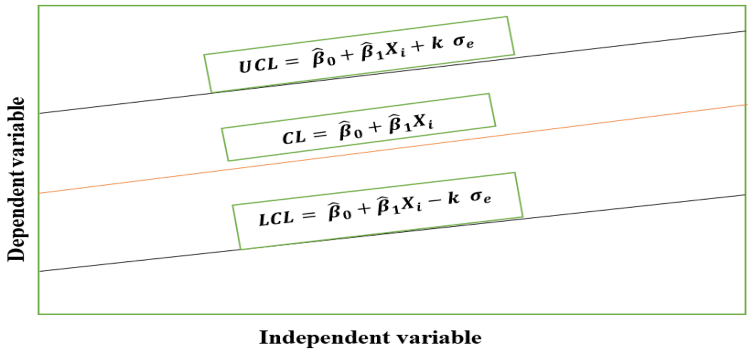

Fig. (2): Regression Control Chart (Aslanand Gunhan, 2014)

3rd: Methods of Normality:

In this research, to analyze data, normality tests were used to help determine if the data is close enough to a normal distribution to use parametric statistical methods. The methods of normality tests used graphical methods: First using Box plots, it can indicate symmetry and the presence of outliers, which can be suggestive of non-normality. Second, using Q-Q plots (Quantile-Quantile plots), it compares the quantities of the data against the quantities of a normal distribution. If the data is normal, the points will fall close to a straight line.

4th: Application Part

The data used in this study is to Building regression control for delectation outliers of analysis from Garota flour factory in Erbil city. The analyze of the data depends on the computer programs IBM Statistic SPSS (Version 28 and Statgraphics 19). The variables to be normalized and two variables were used the variable X: number of working hours at Garota factory is mean responsible variable and Y product of flour. The first part of analyze the data delectation of outliers in the data before analyzing, Fig3.a is shown there is no value of outliers of number of working hours at factory.

Fig3.b is shown there is no value of outliers of product in processing at factory. The real data in Appendix 1. `

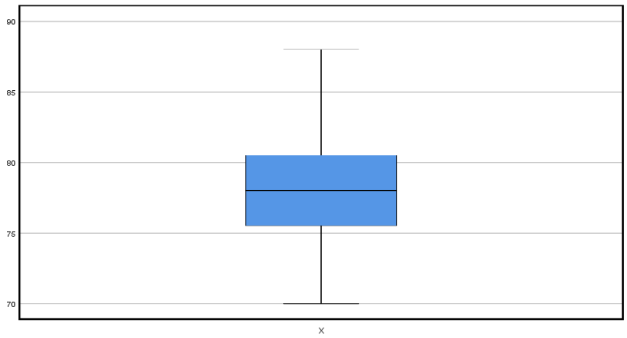

Fig.3 (a): Box Plot of number of working hours at Factory

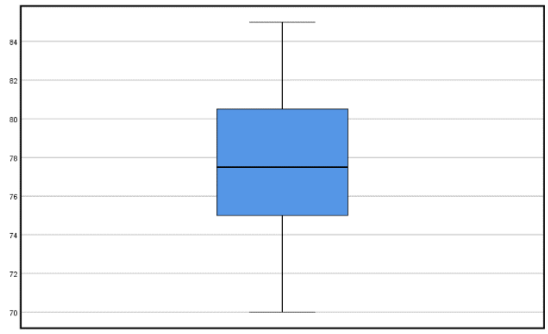

Fig.3 (b): Box Plot of product at factory

Source: output result by the researchers

Table (1): Tests Normality of Variable X

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	**Df**	**Sig.**	**Statistic**	**df**	**Sig.**
X	.078	100	.137	.976	100	.066

Source: output result by the researchers

The second part of analysis data it can be in Table1 below to Determine if the data supports the null hypothesis H0. is distributed according to a normal distribution and using the programs (SPSS) using the normal test and the value of P- value is equal to 0.137 in Kolmogorov-Smirnov, P- value is equal to 0.066 in Shapiro-Wilk also the significance threshold $\alpha = 0.05$ since the null hypothesis is true for this data because the P-value is greater than α. i.e., the data is a normal distribution. This aspect includes an applied study on the use of charts of this data is distributed according to a normal distribution control for the production process, shown in Tabel1 the Fig.4 (a)and(b) are shown a normal distribution of number of working hours at factory, the flour is in control for the production process.

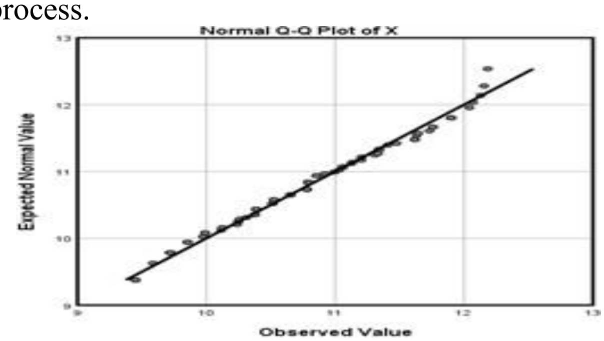

Fig.4(a): Normality Test of number working hours at factory

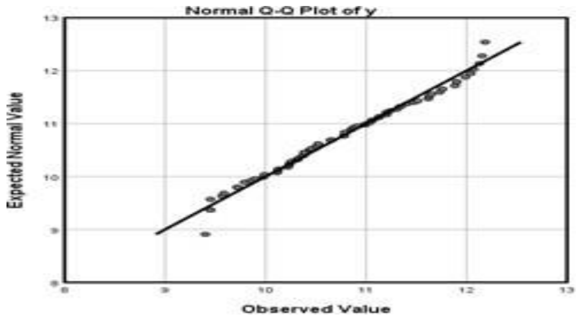

Fig.4(b): Normality Test of produced flour at factory

Source: output result by the researchers

In Table2 below to Determine if the data supports the null hypothesis H0. is distributed according to a normal distribution the value of P- value in Kolmogorov-Smirnov is equal to 0.09, P- value is equal to 0.193 in Shapiro-Wilk. The P- value in is greater than α. i.e., the data is a normal distribution. This aspect includes an applied study on the use of charts is to a normal distribution.

Table (2): Tests Normality of Variable Y

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	Df	Sig.	Statistic	df	Sig.
Y	.082	100	.090	.982	100	.193

Source: output result by the researchers

Depending on testing of normality the null hypothesis H0. is distributed according to a normal distribution, the data is a normal distribution. know estimating parameter by using simple linear regression, it can be shown in Table2.

Table (3): Result of Regression Analysis

Parameter	Estimate	Std. Error	T-Value	P-Value
B0	76.755	7.847	9.781	0.000
B1	0.090	0.842	0.107	0.001

Source: output result by the researchers

In the third part is depend of the parameters, SE, T-Value of the regression model are all significant, the value of B0 is positive meaning that whenever the value of x1 increases by one unit, it leads to an increase in y by an amount, i.e. there is a direct relationship between the two variables, the number of working hours increases, it leads to an increase in wheat production.as it can be seen in Table3 and the R^2 is equal to 87.43% It means that the explanatory variable explains a good percentage statistically, leaving 87.43% to other factors. In other words, if we want to predict based on the model, the results will be accurate with a probability of 12.57. In addition, the intercept and slope parameters are significant because the p-value is less than the significance 0.05. $\sigma_e =$

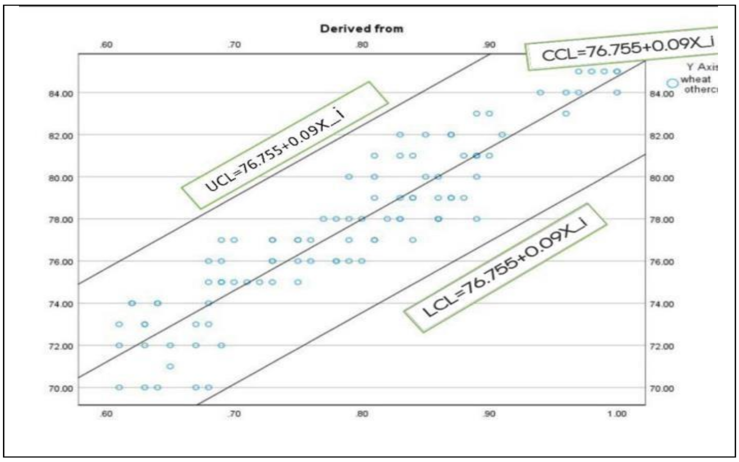

Fig.5: Regression Control Chart of Garota flour factory

$\sqrt{MS_{Residual\ Error}} = \sqrt{13.412} = 3.662$. The target equation is the linear equation model $CCL = 76.755 + 0.09X_i$, the UCL =76.755+0.090Xi +2(3.662) and $LCL = 76.755 + 0.09X_i - 2(3.662)$ Table 4 is shown the of Result of ANOVA

Table (4): Result of one-way ANOVA

Model	Sum of Squares	Df	Mean Square	F	Sig.
Regression	189.779	1	189.779	14.150	.000
Residual	1314.411	98	13.412		
Total	1504.190	99			

Source: output result by the researchers

In the fourth part is drawing the regression control chart it can be shown in Fig5 all of the observation in chart is distributed according to control for the production process. The best target equation is the linear equation model $Y_i$of regression chart, upper control limits and lower control limits. It was shown by using regression chart this method of prediction is better.

Fig6 is the best line regression model can be depended to forecast in our years in future.

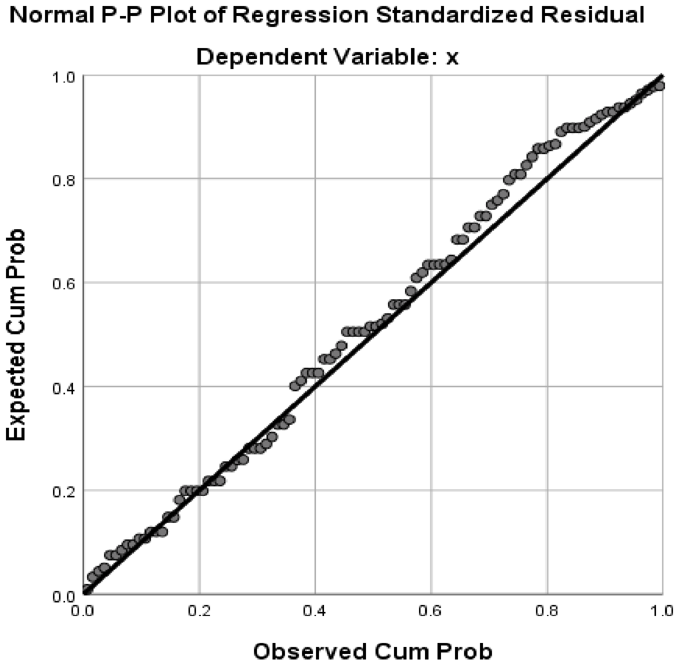

Fig (6): the line regression of factory

Source: output result by the researchers

Table 5 is the Proposed Regression Model's Future Values can depend at product of the Garota flour factory in Erbil to use the hours to work.

Table (5): The Proposed Regression Model's Future Values

Month	2025			2026			2027			2028		
	LCL	Yi^	UCL	LCL	Yi^	UCL	LCL	Yi^	UCL	LCL	Yi^	UCL
1	70.241	77.565	84.889	70.327	77.651	84.975	70.273	77.597	84.921	70.273	77.597	84.921
2	70.251	77.575	84.899	70.305	77.629	84.953	70.295	77.619	84.943	70.262	77.586	84.91
3	70.241	77.565	84.889	70.316	77.64	84.964	70.262	77.586	84.91	70.241	77.565	84.889
4	70.23	77.554	84.878	70.327	77.651	84.975	70.273	77.597	84.921	70.251	77.575	84.899
5	70.219	77.543	84.867	70.197	77.521	84.845	70.219	77.543	84.867	70.284	77.608	84.932
6	70.349	77.673	84.997	70.208	77.532	84.856	70.241	77.565	84.889	70.208	77.532	84.856
7	70.349	77.673	84.997	70.219	77.543	84.867	70.284	77.608	84.932	70.187	77.511	84.835
8	70.338	77.662	84.986	70.208	77.532	84.856	70.241	77.565	84.889	70.316	77.64	84.964
9	70.187	77.511	84.835	70.187	77.511	84.835	70.187	77.511	84.835	70.262	77.586	84.91
10	70.241	77.565	84.889	70.338	77.662	84.986	70.327	77.651	84.975	70.273	77.597	84.921
11	70.284	77.608	84.932	70.349	77.673	84.997	70.349	77.673	84.997	70.338	77.662	84.986
12	70.295	77.619	84.943	70.219	77.543	84.867	70.284	77.608	84.932	70.284	77.608	84.932

Source: output result by the researchers

Conclusion

1. The regression control chart is a useful tool for monitoring processes and forecasting processes when the output quality is important.
2. The control regression control chart provides the factory management with the number of hours needed to produce the specified quantities of flour by forecasting using the regression line equation. This will help the management to schedule working hours.
3. The regression control chart cannot be built if there are outliers, meaning that the data is normally distributed, the regression model is significant, and the regression parameters are statistically significant.

Recommendations

1. This is very important to create research about regression control charts to compare between traditional control charts and the linear regression model.
2. Depended on the computer programs IBM Statistic SPSS to select the outliers in Boxplot to analyze data.
3. Use difference method like Anderson-Darling to know that the data is distributed normally.

References

1- Al-Mashhadani, N. A. (2015). Introduction to Statistical Control of Quality, Al Yamamah Office for Printing and Publishing, Baghdad - Bab Al-Mu'adham, Iraq.
2- Al-Taie, A. E, Al-Hamdani, Kh. A. (2018). Using the Regression Control Chart: An Applied Study on the Production of Badush Cement plant", Iraqi Journal of Statistical Sciences, Vol. 15, No. 27. Johannes
3- Aslan, D.K.; Gunhan, M.B(2014). Regression Control Chart for Auto correlated Processes, Int. J. Industrial and Systems Engineering, Vol. 16, No. 2.
4- Baris O.; Aslan, D.K. (2011). Regression Control Chart for Determination of Young's Modulus: A Case Study, Scientific Research and Essays, Vol. 6, No. 30.
5- Besterfield, D.H. (2004). Quality Control, 7thEdition, Prentice-Hall Inc. New York, USA.
6- Biesterfield, D.H. (1979). Quality Control, Prentice-Hall Inc. New York, USA.
7- Cheng, J. (2007). Six Sigma and TQM in Taiwan: An empirical study. Quality Management Journal, Vol.14, No.2. https://www.slideshare.net/slideshow/control-charts-51167666/51167666
8- Ghosh, B.K., Reynolds Jr, M.R. and Yer, V.H. (1981). Shewhart X-Charts with Estimated Process Variance. Communications in Statistics-Theory and Methods, 10(18), pp.1797-1822.
9- Hasan,S.k.(2018); Use quality control panels to control the ratio of calcium in distilled water ,Al Anbar University journal of Economic and Administration Sciences Vol.10,No.21.
10-Kadir, D.H., Saleh, D.M. and Jamil, D.I. (2024). A Recently Formulated Individual Control Chart Designed for Quality Control Applications within the Health-care Domain. Cihan University-Erbil Scientific Journal, 8(2), pp.9-18.
11-Montgomery, D. C. (2009). Introduction to Statistical Quality Control, 7th ed., Wiley, New York.
12-Montgoomery, D.C. (2005): Introduction to Statistical Quality Control, 5thEdition, John Wiley Sons Inc. New York, USA