4th International Conference on Industrial Engineering and Industrial Management
XIV Congreso de Ingeniería de Organización
Donostia- San Sebastián , September 8th -10th 2010

Implementation of Statistic Process Control in a Painting Sector of a Automotive Manufacturer

J. Cordeiro[1], J. G. Requeijo[2]

1 Dpto. Departamento de Engenharia Mecânica e Industrial, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516, Caparica, Portugal. joanacordeiro@gmail.com
2 UNIDEMI, Dpto. de Engenharia Mecânica e Industrial, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516, Caparica, Portugal. jfgr@fct.unl.pt

Abstract

The continuous improvement on quality of products and processes is a constant concern at organizations, as a response to growing competition and demands of the market. The implementation of statistical techniques adjusted to different situations is one way to achieve this goal. The application of traditional control charts requires, among other, that collected data are independent. However this is not always assured, reflecting a drastic increase of false alarms. This article presents a methodology for application of traditional univariate control charts, based on residuals, when data exhibit significant autocorrelation. Implementation took place in the painting process of an automotive company.

Keywords: Statistic Process Control; Traditional Shewhart Control Charts; Autocorrelated Data; *ARIMA* Models

1. Introduction

The seek for process improvement in industrial organizations implies a better understanding of the nature, interpretation and modeling of variability. Statistical Process Control (*SPC*) is a useful tool in the detection of unusual ways of variation, allowing acting on the same. The control charts, a technique for real-time monitoring of the process, is an appropriate instrument for this purpose.

The univariate control charts were introduced by Walter A. Shewhart, and are currently adopted and implemented in large-scale industrial processes, both discrete and continuous. One of the conditions for the application of these techniques is the existence of data independence, not always found in an industrial environment. If data correlation is disregarded, erroneous conclusions can be taken and derail the identification of statistical control and deviations. According to Gilbert *et al*. (1997), it is imperative to assess if the behavior of autocorrelated data is natural or indisputable to the process. To overcome such adversity, several approaches have been developed regarding the application of univariate control charts to data coming from autocorrelated processes. These approaches are divided into two distinct lines of investigation, a first approach which implies the determination of a mathematical model that best fit the autocorrelated data, and a second without of a mathematical model adjustment.

The statistical control of processes with autocorrelated data has been approached by several researchers, such as Alwan and Roberts (1988), Montgomery and Mastrangelo (1991), Ross and Harris (1991), Maragah and Woodall (1992), Wardell *et al*. (1992), Vander Weil (1996), Zhang (1998), Wieringa (1999), Woodall (2000) and Jay *et al*. (2008), among others.

This paper aims the investigation of the performance and adequacy of univariate traditional control charts to data that exhibit autocorrelation and develops a methodology that allows the application of most appropriate techniques, using the process modeling through *ARIMA* models. The study was performed in a Portuguese automotive industry based on real data.

2. Traditional SPC

For a proper implementation of traditional Shewhart control charts, the sampling must be carried out at appropriate time intervals and sufficient times, so that the collected observations can show the process behavior and maximize the chance of variation between samples; samples should be homogeneous, and it is expected that the units have been produced consecutively and similarly; data, of the characteristic to be controlled, are considered independent and identically distributed according to a Normal distribution with mean μ and variance σ^2; the distance between control limits and centre line of the several charts is ±3 standard deviations of the distribution of sampling statistic to be controlled, which corresponds to a significance level of 0.27%.

There are several references, such as Quesenberry (1997) and Pereira and Requeijo (2008), which consider that the procedure for construction of control charts can be distinguished in, at least, two phases. A first phase (Phase I), where the interest lies in concluding if the data from the past come from a controlled process; a second phase (Phase II), where the control charts are used to monitor the process in real time. It is important to note that Phase II should only be initiated when Phase I of the process is under statistical control, and when the capability to produce in accordance with the required technical specification is verified.

2.1. Phase I

In this phase, named as preliminary phase or retrospective, the upper control limit (*UCL*), lower control limit (*LCL*), center line (*CL*) and the process parameters are estimated. These estimations are made based on collected data, and on the equations presented in Table 1.

Table 1. Control Limits of Shewhart Control Charts for Phase I and estimators of process parameters

Chart	*UCL*	*CL*	*LCL*	$\hat{\mu}$	$\hat{\sigma}$
X (individual observations)	$\overline{X} + 3\left(\overline{MR}/d_2\right)$	$\overline{X}$	$\overline{X} - 3\left(\overline{MR}/d_2\right)$	$\overline{X}$	
$\overline{X}$ (mean)	$\overline{\overline{X}} + A_2\overline{R}$ or $\overline{\overline{X}} + A_3\overline{S}$	$\overline{\overline{X}}$	$\overline{\overline{X}} - A_2\overline{R}$ or $\overline{\overline{X}} - A_3\overline{S}$	$\overline{\overline{X}}$	
R (range)	$D_4\overline{R}$	$\overline{R}$	$D_3\overline{R}$		$\overline{R}/d_2$
S (standard deviation)	$B_4\overline{S}$	$\overline{S}$	$B_3\overline{S}$		$\overline{S}/c_4$
MR (moving range)	$D_4\overline{MR}$	$\overline{MR}$	$D_3\overline{MR}$		$\overline{MR}/d_2$

The parameters (average and standard deviation) are estimated and subsequently the process capability is analyzed. Usually the capability indexes used for Normally distributed data are, $C_p = (LSE - LIE)/(6\sigma)$ and $C_{pk} = \min\left((C_{pk})_I, (C_{pk})_S\right)$, where $(C_{pk})_I = (\mu - LIE)/(3\sigma)$ and $(C_{pk})_S = (LSE - \mu)/(3\sigma)$. Usually the process is considered capable, when C_p and C_{pk} are simultaneously higher than a reference value k (k = 1.33 for bilateral technical specifications and k = 1.25 for unilateral technical specifications).

As mentioned, it is assumed that the data of the characteristic to be controlled are independent and normally distributed. To study the normal distribution, it is suggested the application of the Kolmogorov-Smirnov test and to study the independence of data the application of Autocorrelation Function (*ACF*) and Partial Autocorrelation Function (*PACF*).

After checking the stability of the process, if it does not possess the capability to produce at the technical specification required, corrective actions should be taken, leading to the collection of new data and repeat the Phase I of the *SPC*.

2.2. Phase II

After the verification of stability, parameters estimation, and process capability analysis, the next phase is the monitoring, a procedure commonly referred to as Phase II of *SPC*. The values of the center line and control limits are estimated based on process parameters estimation undertaken in Phase I (equations presented in Table 2), being similar to Phase I if no change of sample size exists.

Table 2. Limits Shewhart Control Charts for Phase II

Chart	*UCL*	*CL*	*LCL*
X (individual observations)	$\mu+3\sigma$	μ	$\mu-3\sigma$
$\overline{X}$ (mean)	$\mu+A\sigma$	μ	$\mu-A\sigma$
R (range)	$D_2\sigma$	$d_2\sigma$	$D_1\sigma$
S (standard deviation)	$B_6\sigma$	$c_4\sigma$	$B_5\sigma$
MR (moving range)	$D_2\sigma$	$d_2\sigma$	$D_1\sigma$

At this stage it is assumed that the process is statistically controlled and the process parameters are known, so when a special cause is present the reason for its occurrence should be investigated and corrective measures taken to proceed to its elimination.

3. SPC for autocorrelated data

As previously mentioned, the use of Shewhart control charts is based on the assumption of statistical independence of the data. When this principle is violated, there are two distinct approaches; a first approach which implies the determination of a mathematical model that best fit the autocorrelated data, and a second without of a mathematical model adjustment.

The first approach involves the study of available data and verifies the type of autocorrelation present in the process, followed by its modeling. In this article we will follow this approach, modeling the process using the *ARIMA* models. After the determination of residuals, or prediction errors, traditional control charts are applied to these variables, following the methodology proposed for the usual traditional *SPC* referred to in section 2.

For the second approach, it is important to refer studies of Montgomery and Mastrangelo (1991) and Montgomery and Mastrangelo (1995). They propose the use of statistical *EWMA* (exponentially weighted moving average) for data from positively autocorrelated processes, as well as the *MCEWMA* chart ((*EWMA* Moving Center Line), which allows simultaneously checking the process control state and monitoring its behavior. Zhang (1998) presents the *EWMAST* chart (*EWMA* for stationary processes), which is nothing more than a change in the control limits of *EWMA* control chart through the autocorrelation function ρ_k.

3.1. Shewhart charts for autocorrelated data

Modeling based on *ARIMA* models, plays a filter role that allows the elimination of the existing process autocorrelation, making residuals independent and normally distributed. With these residuals, Shewhart control charts are designed for Phase I of *SPC*. Later, when the process is under statistical control, forecasts will be carried to subsequent periods and calculated the prediction errors. These prediction errors represent the set of data to be monitored in Phase II of *SPC*.

3.1.1. Shewhart charts for residuals control

Like *SPC* for independent data, the techniques to be implemented in Phase I are the traditional Shewhart charts applied to residuals (determined with the process modeling). In Table 3 the equations that allow the estimation of control limits and the standard deviation of the residuals, σ_ε, are presented.

Table 3. Control Limits for Phase I of control charts based on residuals and estimators σ_ε

Chart	*UCL*	*CL*	*LCL*	$\hat{\sigma}_\varepsilon$
e (residuals)	$3\left(\overline{MR}/d_2\right)$	0	$-3\left(\overline{MR}/d_2\right)$	
$\bar{e}$ (mean)	$A_2\bar{R}$ or $A_3\bar{S}$	0	$-A_2\bar{R}$ or $-A_3\bar{S}$	
R (range)	$D_4\bar{R}$	$\bar{R}$	$D_3\bar{R}$	$\bar{R}/d_2$
S (standard deviation)	$B_4\bar{S}$	$\bar{S}$	$B_3\bar{S}$	$\bar{S}/c_4$
MR (moving range)	$D_4\overline{MR}$	$\overline{MR}$	$D_3\overline{MR}$	$\overline{MR}/d_2$

The distribution of residuals has an expected zero mean value and variance σ_ε^2. Based on the adjusted *ARIMA* model (*AR* (*p*), *MA* (*q*) or *ARMA* (*p*, *q*)), the location (μ) and dispersion (σ) process parameters are estimated. These models will be described later on this paper. It is assumed that the residuals, concerning to the characteristic under study, are independent and Normally distributed. It is suggested that verification of the Normality of the residuals is checked using the Kolmogorov-Smirnov test, and the independence by the *ACF* and *PACF*.

When a special cause of variation is found, the point should not be eliminated but replaced by the expected value for that moment, setting back the *ARIMA* model, calculate the new residuals, and review the control charts. If there are many points out of control, it is necessary to investigate the cause(s) that led to this situation and proceed to the necessary corrective actions. During the checking of the stability of the process, if it does not reveal capability to produce according to technical specification, corrective actions should be taken.

One of the objectives from Phase I of the *SPC* is to estimate properly the process parameters. When the quality characteristic *X* shows significant autocorrelation, this estimation is performed based on the parameters of the *ARIMA* model. The estimators of process mean and variance for the different models are given by the equations on Table 4. In the equations of Table 4 it is considered σ_ε^2 the variance of residuals, ξ the parameter that determines the process average, ϕ_j the order parameter *j* of *AR* or *ARMA* model, θ_j the order parameter *j* *MA* or *ARMA* model, ρ_j the correlation coefficient of lag *j* and γ_j auto-covariance lag *j*. The study of the process capability is performed using the traditional capability indexes of process C_p and C_{pk}, previously defined.

Table 4. Estimators of mean and variance of process, on the quality characteristic *X*

Model	Mean	Variance
AR	$E(X) = \mu = \frac{\xi}{1-\sum_{j=1}^{p}\phi_j}$	$Var(X) = \frac{\sigma_\varepsilon^2}{1-\sum_{j=1}^{p}\rho_j \cdot \phi_j}$
MA	$E(X) = \mu$	$Var(X) = \sigma_\varepsilon^2 \cdot \sum_{j=0}^{q}\theta_j^2 \;,\; \theta_0 = 1$
ARMA	$E(X) = \mu = \frac{\xi}{1-\sum_{j=1}^{p}\phi_j}$	$Var(X) = \sum_{j=1}^{p}\phi_j \cdot \gamma_j - \theta_1 \cdot \gamma_{X\varepsilon}(-1) - \ldots - \theta_q \cdot \gamma_{X\varepsilon}(-q) + \sigma_\varepsilon^2$

The monitoring of future data, for autocorrelated processes, is performed through the application of Shewhart control charts based on the prediction errors. The distribution of prediction errors presents an expected mean value of zero and a variance of $Var(e_t(\tau)) = \sigma_\varepsilon^2\left(1+\sum_{j=1}^{\tau-1}\psi_j^2\right)$. The prediction errors are given by $e_\tau(\tau) = X_{T+\tau}(\tau) - \hat{X}_{T+\tau}(\tau)$ and the limits of control charts of prediction errors are defined by expressions that are presented in Table 5, where $\sigma_{ep} = \sqrt{Var(e_\tau(\tau))}$.

Table 5. Control Limits for Phase II of control charts based on prediction errors

Chart	*UCL*	*CL*	*LCL*
e (residuals)	$3\sigma_{ep}$	0	$-3\sigma_{ep}$
$\bar{e}$ (mean)	$A\sigma_{ep}$	0	$-A\sigma_{ep}$
R (range)	$D_2\sigma_{ep}$	$d_2\sigma_{ep}$	$D_1\sigma_{ep}$
S (standard deviation)	$B_6\sigma_{ep}$	$c_4\sigma_{ep}$	$B_5\sigma_{ep}$
MR (moving range)	$D_2\sigma_{ep}$	$d_2\sigma_{ep}$	$D_1\sigma_{ep}$

4. *ARIMA* Methodology of Box and Jenkins

The aim of the methodology of Box and Jenkins is to determine and adjust the better *ARIMA* mathematical model to the collected process observations, in order to eliminate the existing autocorrelation and obtain a good prediction for each observation. To model a process by using this methodology, it is necessary to determine the *ARIMA* (*p*, *d*, *q*) model that best fits the data, comparing the estimated autocorrelation function (*EACF*) with the theoretical autocorrelation function (*ACF*) and the estimated partial autocorrelation function (*EPACF*) with the theoretical partial autocorrelation function (*PACF*). The modeling based on the *ARIMA* methodology follows certain steps, namely: identification, estimation, verification, prediction. In stationary processes, *ACF* and *PACF* of process *AR*(*p*), *MA*(*q*) and *ARMA*(*p*,*q*), have distinct characteristics. Table 6 gives the different succession series.

Table 6. Succession series

Process	***ACF***	***PACF***
AR(*p*)	Exponential decrease from a certain lag order	Presents significant peaks through lag (*p*) which indicate the order of the model
MA(*q*)	Presents significant peaks through lag (*q*) which indicate the order of the model	Exponential decrease from a certain lag order
ARMA(*p*,*q*)	Exponentially decrease from a certain order lag, positively or negatively, or switching between them	

Thereafter, the identified model parameters are estimated: the parameters ϕ if the behavior of the process is autoregressive; the parameters θ for moving average and variance of the error σ_{ε}^2. When a satisfactory model is obtained, it is possible to predict the potential future values of the characteristic under study and then determine the forecasting errors.

5. Study Case

The presented study case includes the study of the quality characteristic "Total Thickness", and corresponds to the paint industry process of a Portuguese automotive manufacturer. This study aimed the process control, considered of great relevance for the company. The data collected during production showed the existence of significant autocorrelation, and two different approaches have been developed and their performance compared. The first approach consisted in the implementation of the methodology suggested by the authors, using the process modeling and application of Shewhart charts of residuals. The second approach was to apply the traditional charts directly to the collected data without the concern of studying a possible significant autocorrelation.

5.1. Verification of data autocorrelation

To study the data autocorrelation, 93 observations related to Total Thickness are collected. The verification of autocorrelation is performed using the "Statistica" software, calculating the *EACF* and *EPACF* for the 93 collected observations. The analysis of Figure 1 reveals the existence of significant autocorrelation of data, since the coefficient of correlation estimated for the lag 1 (0.595) does not belong to its confidence interval. Comparing the *EACF* and *EPACF* with the *ACF* and *PACF* described in Table 6, it was found that the process can be modeled using a standard *AR*(1).

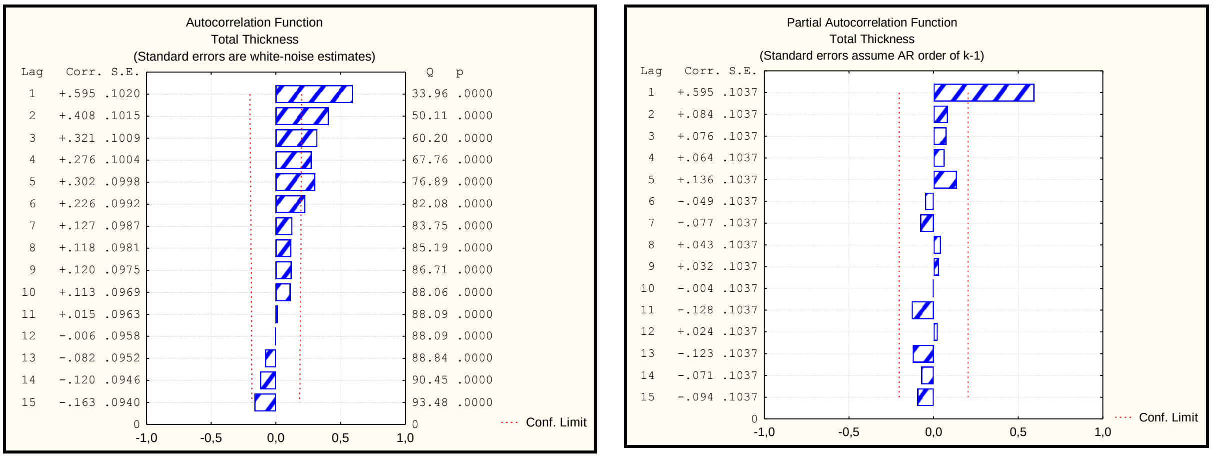

Figure 1. *FACE* and *FACPE* for Total Thickness characteristic

To estimate the parameters of the model *AR*(1), the software "Statistica" is used, obtaining the values $\hat{\mu} = 82.038$ and $\hat{\phi}_1 = 0.59474$. After the process modeling, the residuals independence was checked with *EACF* and *EPACF* of residuals, presented in Figure 2. The independence of the residuals is found.

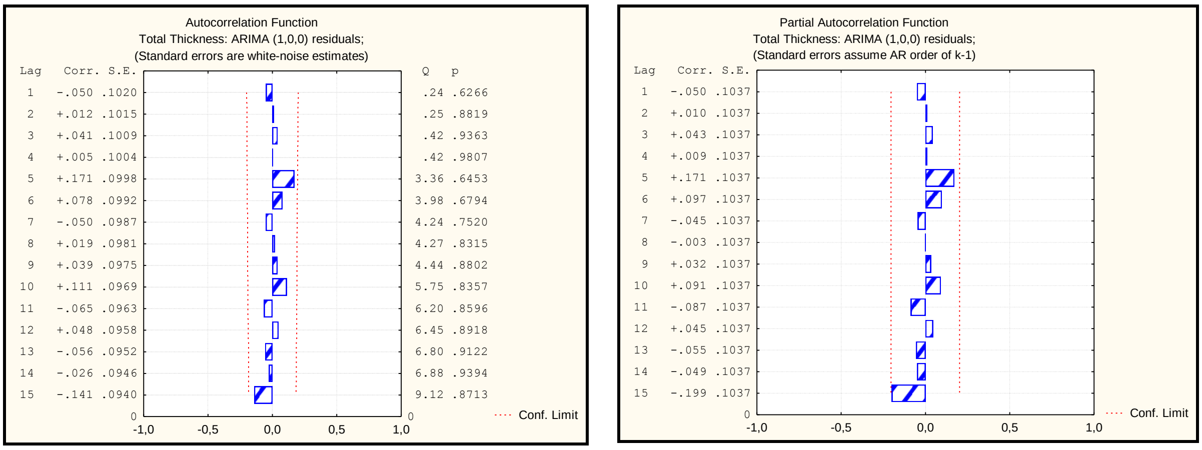

Figure 2. *FACE* and *FACPE* for residuals of Total Thickness characteristic

5.2. SPC Application (Phase I)

5.2.1. Control charts applied to residuals

For the Shewhart control charts with autocorrelated data, designed to be applied when the observations violate the assumption of independence, the appropriate methodology is the following: with the 93 values used for the autocorrelation study, the control charts applied to residuals are drawn (Figure 3).

During the construction of the charts, it was noticed that point nr.10 was a special cause of variation, since it exceeded the value of the *UCL* chart *MR*; the responsibility of this situation ($MR_{10} > LSC_{MR}$) lays on the value X_9, reason why this value was replaced by its expected value ($\hat{X}_9 = \xi + \phi_1 X_8 = 82.038\times\left(-0.59474\right)+0.59474\times 77.2 = 79.2$), the *ARIMA* model is then fitted and the revised charts of residuals are built (Figure 3). By examining this figure it is concluded that there are no special causes of variation and the process is under statistical control. It is presented in Table 7 the new values of model parameters.

Table 7. Parameters of the model AR (1) for Total Thickness characteristic

Model (1,0,0) MS Residual = 2.2486						
	Parameter	**Standard Deviation**	t_0	***p-value***	**Confidence Interval 95%**	
					Lower Limit	**Upper Limit**
Constant	82.064	0.3694	222.2	0.000000	81.3301	82.7974
ϕ_1	0.58735	0.08532	6.88	0.000000	0.41787	0.75683

The Normality of the residuals is checked using the Kolmogorov-Smirnov test ($d = 0.07447$; $D_{crítico} = 0.886/\sqrt{N} = 0.0919$ to $\alpha = 5\%$; $d < D_{crit}$). It is presented in Figure 4 the histogram of 93 values with the Kolmogorov-Smirnov value, and the *p*-value of Chi-Square test.

Checked the stability of the process, the parameters of Total Thickness characteristic are estimated, based on the model *AR*(1), which are presented in Table 8.

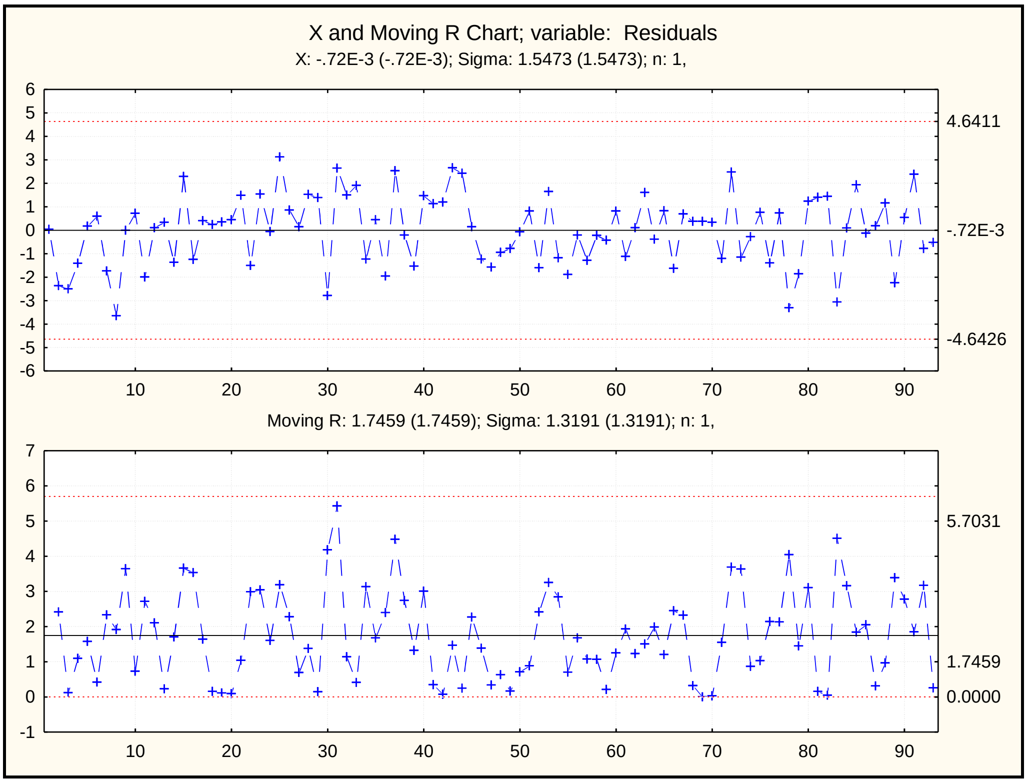

Figure 3. Control chart *e-MR* of revised residuals of the characteristic Total Thickness

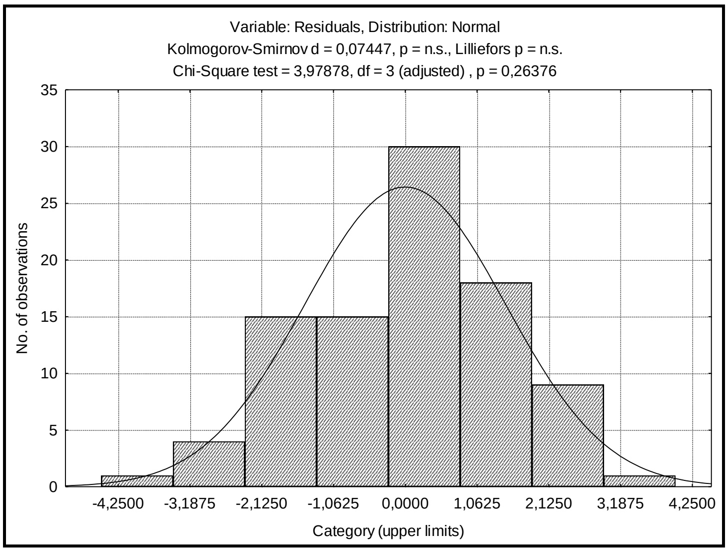

Figure 4. Check of the Normality of residuals of Total Thickness characteristic

Table 8. Parameters for Thickness Total characteristic

	Model			Control Chart	Process	
Parameter	$\hat{\mu}$	$\hat{\phi}_1$	$\hat{\rho}_1$	$\hat{\sigma}_\varepsilon$	$\hat{\mu}$	$\hat{\sigma}$
Estimate	82.064	0.58735	0.587	1.5473	82.064	1.9115

After the parameters estimation it is possible to study the process capability. Once the specification is unilateral ($LSL = 70$), the study of capability is carried out only based on the C_{pk} ($C_{pk} = 2.10$). It is possible to conclude that the process is able to produce at the required technical specification.

5.2.2. Control chart for data without modelation

When the Shewhart control charts are constructed directly from the original autocorrelated data, the limits of traditional control, increases the chance of false alarms. This is illustrated in Figure 5, where it is presented the *X-MR* control charts for 93 data of Total Thickness characteristic. The analysis of these charts indicates erroneously the existence of many special causes of variation in the process (in *X* chart, the observations 3, 8, 9, 43, 44, of whom only in instant 9 a special cause of variation exists and the remainder are false alarms).

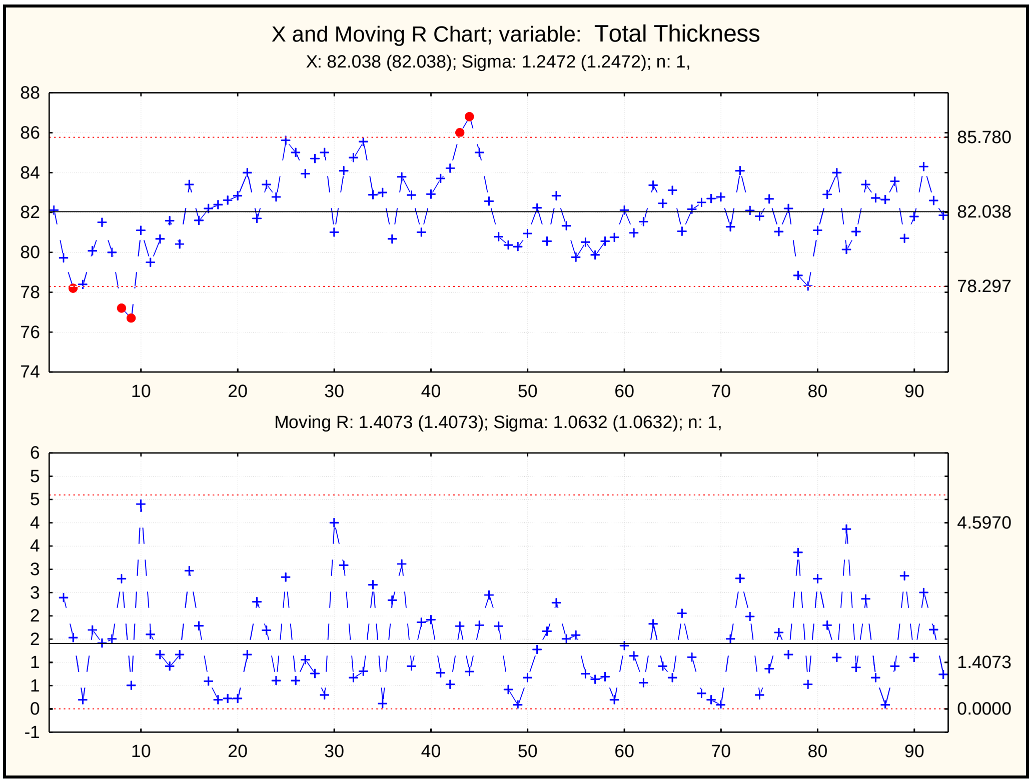

Figure 5. Control chart *X-MR* of original data of the characteristic Total Thickness

5.2.1. Comparison of the study with and without modelation

Comparing both methods, and analyzing Figure 3 (residuals chart after modelation) and Figure 5 (charts applied to original data without modelation), it is quite severe not to consider the existing process autocorrelation. The indication given with the application of these statistical techniques (Figure 5) may incur in severe analysis errors which are dangers for the company, because it will impure resources to investigate some anomalies, when the process is actually stabilized, without special causes of variation.

6. Conclusions

The development of a statistical control methodology appropriated to the existence of significant process autocorrelation, reveals a crucial importance, since it avoids possible analysis errors. If the autocorrelation is not considered, some mistakes may occur, namely: 1) consider a stable process, when several special causes of variation are present, which corresponds to a non stable process; 2) consider a process out of statistical control, when it is actually stable (increase of false alarms); 3) incorrect estimate of the process parameters; 4) proceed to rough and/or incorrect analysis of the process capability; 5) loss of resources by a unnecessary intervention on the productive process, in order to solve problems which are not real (false alarms).

Bibliography

Alwan, L.; Roberts, H. (1988). Time-Series Modeling for Statistical Process Control. Journal of Business & Economic Statistics, Vol. 6, pp. 87-95.

Gilbert, K. C.; Kirby, K.; Hild, C. R. (1997). Charting Autocorrelated Data: Guidelines for Practitioners. Quality Engineering, Vol. 9 , pp. 367-382.

Harris, T. J.; Ross, W. H. (1991). Statistical Process Control Procedures for Correlated Observations. Canadian Journal of Chemical Engineering, Vol. 69, pp. 48-57.

Junior, F. J.; Siedel, E. J.; Lopes, L. F. (2008). Comparação entre métodos utilizados no tratamento de dados autocorrelacionados no controle estatístico do processo. XXVIII encontro nacional de engenharia de produção. Rio de Janeiro, Brasil.

Maragah, H. D.; Woodall, H. W. (1992). The effect of autocorrelation on the retrospective X-chart. Journal of Statistical Computation and Simulation, Vol. 40, pp.29–42.

Mastrangelo, C. M.; Montgomery, D. C. (1995). SPC with Correlated Observations for the Chemical and Process Industries. Quality and Reliability Engineering International, Vol. 11, pp. 79-89.

Montgomery, D. (2008b). Introduction to Statistical Quality Control. 6th ed.. New York / Jonh Wiley & Sons.

Montgomery, D. C.; Mastrangelo, C. M. (1991). Some Statistical Process Control Methods for Autocorrelated Data. Journal of Quality Technology, Vol. 23, pp. 179-193.

Pereira, Z. L.; Requeijo, J. G. (2008). Qualidade: Planeamento e Controlo Estatístico de Processos. Lisboa / Prefácio.

Quesenberry, C. P. (1997). Methods for Quality Improvement. New York / John Wiley & Sons.

Vander Weil, S. A. (1996). Monitoring Processes that Wander Using Integrated Moving Average Models. Technometrics, Vol. 38, pp. 139-151.

Wardell, D. G.; Moskowitz, H.; Plante, R. D. (1992). Control Charts in the Presence of Data Correlation. Management Science, Vol. 38, pp. 1084-1105.

Wieringa, J. E. (1999). Statistical Process Control for Serially Correlated Data. Ph.D. Thesis University of Groningen.

Woodall, W. H. (2000). Controversies and Contradictions in Statistical Process Control. Journal of Quality Technology, Vol. 32, pp. 341-350.

Zhang, N. F. (1998). A Statistical Control Chart for Stationary Process Data. Technometrics, Vol. 40, pp. 24-38.