

The Usage of Time Series Control Charts for Financial Process Analysis

Kovářík Martin, Klímek Petr

Abstract

We will deal with financial proceedings of the company using methods of SPC (Statistical Process Control), specifically through time series control charts. The paper will outline the intersection of two disciplines which are econometrics and statistical process control. The theoretical part will discuss the methodology of time series control charts and in the research part there will be this methodology demonstrated in three case studies. The first study will focus on the regulation of simulated financial flows for a company by CUSUM control chart. The second study will involve the regulation of financial flows for a heteroskedastic financial process by EWMA control chart. The last case study of our paper will be devoted to applications of ARIMA, EWMA and CUSUM control charts in the financial data that are sensitive to the mean shifting while calculating the autocorrelation in the data. In this paper, we highlight the versatility of control charts not only in manufacturing but also in managing the financial stability of cash flows.

Keywords: Statistical Process Control, Shewhart's control charts, autocorrelation, EWMA control chart, CUSUM control chart, ARIMA control chart.

1. INTRODUCTION

Traditional SPC schemes can be applied to monitoring the residuals. Subsequent work on this problem can be broadly classified into two themes; those based on time series models and those which are model-free. For the former, three general approaches have been proposed: those which monitor residuals, those based on direct observations, and those based on new statistics. A brief account of these approaches is presented in this chapter. The time series model based approach is easy to understand and effective in some situations. However, it requires identification of an appropriate time series model from a set of initial in-control data. In practice, it may not be easy to establish and may appear to be too complicated to practicing engineers. Hence, the model-free approach has recently attracted much attention. The most popular model-free approach is to form a multivariate statistic from the autocorrelated univariate process, and then monitor it with the corresponding multivariate control chart. (Krieger *et al.*, 1992) used a multivariate CUSUM scheme. (Apley and Tsung, 2002) adapted the T^2 control chart for monitoring univariate autocorrelated process. (Atienza *et al.*, 1997) proposed a Multivariate boxplot- T^2 control chart. (Dyer *et al.*, 2003) adapted the use of the multivariate EWMA control chart for autocorrelated processes. Statistical financial flow proceeding means the cash flow management in company. We can avoid possible loss by the cash flow monitoring. This loss can be caused by nondelivery goods, bad financial investment, etc. Financial analysis should be done once a year. For our example, we will introduce regulation of simulated financial flows for a company by CUSUM control chart (see Case Study No 1). In other example, we will describe EWMA control chart using also monthly values (see Case Study No 2). The end of this paper will be dedicated to the ARIMA, EWMA and CUSUM control charts together with some practical example of autocorrelated data (see Case Study No 3).



2. ELEMENTAR DATA ASSUMPTIONS

This part is mainly focused on crucial problems with the statistical analysis data assumptions. Fundamental assumptions for statistical process regulation can be described as:

- data normality, symmetry,
- constant mean of the process,
- constant variance (standard deviation) of data,
- independence, no autocorrelation in data,
- absence of outliers (Meloun and Militký, 2006).

Most data analysis processes and their conclusions are dependent on some fulfilled conditions. If they are not fulfilled all other calculations of means, confidence intervals, quantiles, statistical tests, Shewhart's charts, capability indices are questionable and not really correct. These calculations usually offer incorrect and inaccurate results and conclusions. Therefore we should be very careful about above mentioned conditions (data normality, symmetry, etc.). Violations of assumptions for the application of regulation by Shewhart's charts in different technologies are displayed in (Meloun and Militký, 2006).

Mentioned conditions should be verified by the help of statistical tests. For example, we can meet data asymmetry by the physical quantities such as strength or viscosity. We can meet strong autocorrelation (dependence) in continuous processes in chemistry, pharmacy, food and metals. Quality of input process material can result into the mean shifting.

Not normally distributed data can be seen in ecologic processes very often. Data are very asymmetric with usually lognormal distribution.

3. LITERATURE REVIEW

3.1 Control Charts CUSUM

The CUSUM control charts are based on the cumulative sums. They were introduced by Page in 1954. Their main advantage is a very quick detection of relatively small shift in the process mean. This detection is significantly quicker than by the Shewhart's control charts.

The sequential sums of deviations from μ_0 are used for the CUSUM control chart construction. If μ_0 is a target value for the population mean and if X_j is a sample mean then the CUSUM control chart is constructed by plotting of variables of the $S_i = \sum_{j=1}^i (X_j - \mu_0)$ type. This process is called a random walk (Harris and Ross, 1991).

3.2 CUSUM – Chart for Individual Values and for Samples Means from Normally Distributed Data

Values of x_i are independent with the same normal distribution $N(\mu, \sigma^2)$ with the known population mean and with the known population standard deviation σ . We assume logical subgroups with the same volume n . Cumulative sum – CUSUM C_n is defined for individual values ($n = 1$) as:

A) on a base of original scale:

$$C_n = \sum_{j=1}^n (x_j - \mu) \quad (1)$$

B) on a base of normal distribution where the mean = 0 and the standard deviation = 1:

$$U_j = \frac{(x_j - \mu)}{\sigma}, S_n = \sum_{j=1}^n U_i \quad (2)$$

The CUSUM C_n is almost the same as CUSUM S_n measured in the units of standard deviation σ . Equation for C_n can be written recurrently (Chandra, 2001): $C_0 = 0$,

$$C_n = C_{n-1} + (x_n - \mu); \text{ and with the same prnciple for } S_n: S_0 = 0, S_n = S_{n-1} + U_n.$$

Suppose that the original distribution of observed variable $N(\mu, \sigma^2)$ changes into $N(\mu + \delta, \sigma^2)$ distribution for integer t (at certain moment). It means that the population mean μ will face a certain shift of δ . It also means that the shift starts at point (m, C_m) and it grows linearly with the slope δ . But the population means shift can be more complicated. The CUSUM control chart can reflect all these changes (Harris and Ross, 1991).

3.3 CUSUM for Sample Means \bar{x}_j

We have considered mainly the individual values until now. Now, we will consider subgroups with m observations and we calculate the sample means from this subgroups. We have to work with the sample mean standard deviation $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{m}}$. A shift of mean Δ will not be measured in the units of σ but in the units of $\sigma_{\bar{x}}$ in this case. We will substitute the individual values of x_i with the sample means \bar{x}_j and the process standard deviation σ with the sample mean standard deviation $\sigma_{\bar{x}}$ in above mentioned formulas. (Lu and Reynolds, 1999)

New Process Mean Estimate

If there is a shift we can estimate a new process mean from the next formula:

$$\hat{\mu} = \begin{cases} \mu_0 + K + \frac{C_1^+}{N^+} & \text{pro } C_1^+ > H \\ \mu_0 - K + \frac{C_1^-}{N^-} & \text{pro } C_1^- < -H \end{cases}, \quad (3)$$

where N^+ and N^- is a number of selected points from a moment (Chambers and Wheeler, 1992), when $C_n^+ = 0$, resp. when $C_n^- = 0$.

Comparison of CUSUM and Shewhart's Control Charts

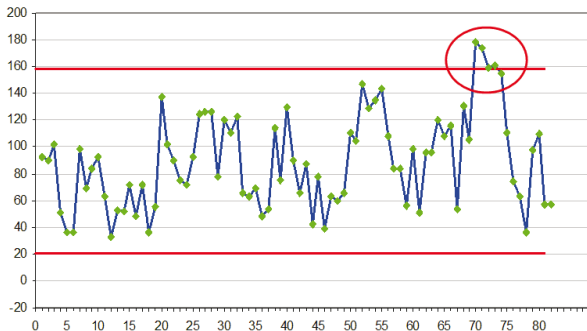


Fig. 1 – Shewhart's Control Chart. Source: QC Expert 2.5C \bar{x}

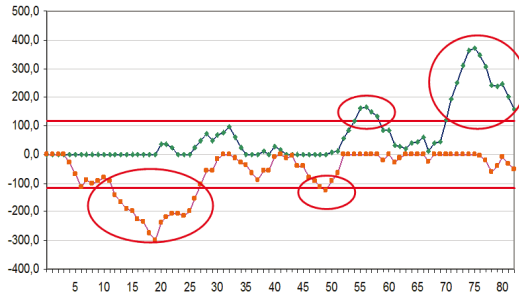


Fig. 2 – Control Chart CUSUM. Source: *QC Expert 2.5C*

This example shows practically sensitivity of the CUSUM control chart in comparison with the Shewhart’s control chart for the sample means. The CUSUM control chart detects process mean deviation towards the lower values (around the subgroup 20 – see Figure 1) while the Shewhart’s control chart does not detect this deviation (Harris and Ross, 1991). It does not detect a shift towards the upper values (around the subgroup 56). It only detects a big shift around the subgroup 70 (see both Figures 1 and 2) (Lu and Reynolds, 1999).

3.4 Dynamic Control Chart EWMA

Dynamic control charts EWMA (Exponentially Weighted Moving Average) are used when the following conditions are fulfilled:

- are not independent, with positive autocorrelation
- mean is not constant, its changes are slow (Montgomery and Friedman, 1989)

A sudden change in mean will only cause a control limits crossing. These dynamic charts provide not only the information about “in control” process but also about the process dynamic development. As we mentioned, we consider only data which are not independent with positive autocorrelation. We will explain it now. If the measured observations are influenced by the previous ones we can say that they are dependent. A special case of this dependence is so called autocorrelation of 1st degree when this dependence is linear. If there is a positive autocorrelation in data then the smaller value follows after the smaller value and the higher value follows after the higher value. Data have tendency to preserve their original values. Process is not stable in a case of negative autocorrelation. If there is a negative autocorrelation in data then the higher value follows after the smaller value and the smaller value follows after the higher value.

Suppose that we measure values x_1, x_2, x_3, \dots for the variable X in the process. Parameter λ (level of “forgetting”) is calculated by trying where the function

$$\sum_{k=1}^n e_k^2$$

is minimal. Number n is equal to number of measured values of regulated variable. It is recommended that n is greater than 50. If the error values of one-step prediction of e_k for the optimal value of parameter λ are not correlated and if they have a normal distribution then the center line CL_k , control limits UCL_k and LCL_k for the dynamic control chart EWMA are calculated from the following formulas (Lu and Reynolds, 1999):



$$\text{CL}_k = \hat{x}_{k-1}, \quad \text{LCL}_k = \hat{x}_{k-1} - \hat{\sigma}_p u_{1-\frac{\alpha}{2}}, \quad \text{UCL}_k = \hat{x}_{k-1} + \hat{\sigma}_p u_{1-\frac{\alpha}{2}}, \quad \hat{\sigma}_p^2 = \frac{1}{n-1} \sum_{k=1}^n e_k^2$$

where $\hat{\sigma}_p^2$ is a standard deviation of e_k estimate, while e_k values are determined for the optimal value of parameter λ (Yourstone and Montgomery, 1991).

3.5 ARIMA Control Chart

Classical Shewhart SPC concept assumes that the measured data are not autocorrelated. Even very low degree of autocorrelation data causes failure of classical Shewhart control charts. Failure has a form of a large number of points outside the regulatory limits in control diagram. This phenomenon is not unique in the case of continuous processes, where the autocorrelation data given by the inertia processes in time (chemical processes, climate processes etc.). Autocorrelation of data becomes increasingly frequent phenomenon in terms of discrete processes, in particular manufacturing processes with short production cycles, high speed production with a high degree of automation of production and also in test and control operations. In these conditions, it is possible to obtain data about each product, with the consequence that the time between measurements (recording) of two consecutive values of the monitored variables is very short. One of the ways to tackle autocorrelated data is the concept of stochastic modeling of time series using autoregressive integrated moving average models, the ARIMA model. The concept of linear stochastic autoregressive models (models AR), moving average (model MA), mixed models (the ARMA models), and ARIMA models, based on Box-Jenkins methodology is seen as a time series realization of stochastic process. Box-Jenkins methodology represents a modern concept analysis of stationary and nonstationary time series based on probability theory. Linear models AR, ARMA and MA are modeling tool for the stationary processes. These models have a characteristic shape of the autocorrelation fiction (Autocorrelation Fuction – ACF) and partial autocorrelation function (Partial Autocorrelation Function – PACF), which are an essential tools for providing information about the stochastic process. ACF and PACF estimates are used to identify the time series model. There are non-stationary processes in practice very often. Non-stationarity can be present due to the mean value changing over time or process variance changing over time. If, after the transformation of nonstationary process variance of “random walk” (so-called integrated process) using differential d -th order is the final process model to describe the stationary ARMA (p, q), the original integrated process is called an autoregressive integrated moving average process of order p, d, q , ie ARIMA (p, d, q) (Noskievičová, 2008).

ARIMA control chart (Autoregressive Integrated Moving Average) is based on the principle of finding a suitable time series model and the use of control chart for residuals the model (deviations from the values actually measured values from calculated values with the model use).

The general shape of the model ARIMA (p, d, q) is such

$$\Phi_p(B) \cdot \nabla^d \cdot x_t = \Theta_q(B) \varepsilon_t, \quad (4)$$

where

$\Phi_p(B) = (1 - \phi_1 B - \phi_2 B^2 \dots - \phi_p B^p)$ is autoregressive polynomial of p -th order,

$\Theta_q(B) = (1 - \theta_1 B - \theta_2 B^2 \dots - \theta_q B^q)$ is moving averages polynomial of q -th order,

(grad) operator backward difference (this element is introduced when the model exhibits non-stationarity of the process),

d difference order,

t time,

B back shift operator $B \cdot x_t = (x_{t-1})$,

$\phi_1, \phi_2, \dots, \phi_p$ parameters of autoregressive model,

$\theta_1, \theta_2, \dots, \theta_q$ parameters of moving averages model,

ε_t this variable is called white noise, which is unpredictable fluctuations in the data, has a normal distribution with mean equal to zero and constant variance, and its values are uncorrelated.

If \hat{x}_t is an estimate of empirical value of x_t calculated with help of right chosen ARIMA model, residuals of this model will be uncorrelated normally distributed random variables.

Most commonly used in applications are ARIMA models. Let us consider the model

$$x_t = \xi + \phi x_{t-1} + \varepsilon_t \quad (5)$$

where ξ a ϕ ($-1 < \phi < 1$) are unknown constants and ε_t is normally distributed and uncorrelated variable with the mean equal to zero and the constant standard deviation σ . This model is called autoregressive model of the first order and is denoted as AR(1). The values of the reference mark of quality, which are mutually shifted of k time periods (x_t and x_{t-k}) have the correlation coefficient ϕ^k . This means that autocorrelation function ACF should fall exponentially. If we expand the previous equation in the form

$$x_t = \xi + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \varepsilon_t \quad (6)$$

we get equation of second order autoregressive model AR(2). Generally, variable x_t is directly dependent on the values preceding x_{t-1} , x_{t-2} , etc. in the autoregressive model AR (p). If we model the dependence of data using the random component ε_t , then we get moving average model MA (q). Moving average model first order has an equation:

$$x_t = \mu + \varepsilon_t - \theta \varepsilon_{t-1} \quad (7)$$

There is some correlation only between two values x_t and x_{t-1} . It can be described as follows: $\rho_t = -\theta / (1 + \theta)$. This corresponds to the shape of the autocorrelation function ACF (Noskievičková, 2008). For the modeling of practical problems is often suitable a model compound containing both an autoregressive component and the moving averages component. This model is generally known as ARMA (p, q) (Hušek, 2007). Model ARMA of the first order, i.e. ARMA (1, 1) is described by the equation:

$$x_t = \xi + \phi x_{t-1} + \varepsilon_t - \theta \varepsilon_{t-1} \quad (8)$$

This model is often suitable for chemical and other continuous processes, where many quality characteristics can be easily modeled by the model AR (1). Measurement errors are described by random component of the model, which we assume to be random and uncorrelated. The ARMA model assumes stationarity process i.e., that the character quality reference values are around a stable mean. But often, in practice there are processes (e.g. in the chemical industry), where the values of monitored variable are “running away”. Then it is convenient to model processes using appropriate model with the operator of backward difference grad, such as the ARIMA model (0, 1, 1), the equation is

$$x_t = x_{t-1} + \varepsilon_t - \theta \varepsilon_{t-1} \quad (9)$$

Models ARIMA are different from Shewhart model ($x_t = \mu + \varepsilon_t$ for $t = 1, 2, \dots$). However, if we put $\varphi = 0$ in the equation $x_t = \zeta + \theta x_{t-1} + \varepsilon_t$ or $\theta = 0$ in the equation $x_t = \mu + \varepsilon_t - \theta \varepsilon_{t-1}$ we get Shewhart model process. Another important step in the use of ARIMA models is the choice of the appropriate SPC control chart. When residuals testing determined that they are not autocorrelated and come from a normal distribution it is possible with them to verify whether or not the process is statistically stable. Because number of observations equals one (original empirical values were detected by each unit) priority have control charts for individual values and moving range. Location of the mean value CL and upper and lower control limits (UCL, LCL) for the ARIMA chart for individual values can be determined from the formula

$$CL = \bar{e} (\cong 0) \tag{10}$$

$$UCL = \bar{e} + \frac{3}{1,128} \bar{R}_k \tag{11}$$

$$LCL = \bar{e} - \frac{3}{1,128} \bar{R}_k \tag{12}$$

Values CL , UCL and LCL can be calculated as follows

$$CL = \bar{R}_k \tag{13}$$

$$UCL = 3,267 \cdot \bar{R}_k \tag{14}$$

$$LCL = 0 \tag{15}$$

To increase the sensitivity of control charts ARIMA is recommended to use two-sided CUSUM control chart with the decision interval $\pm H$ or standard EWMA chart, both applied to the residuals of the model. If we pursue more quality characteristics simultaneously on a single product to multiple, we can apply to ARIMA models residuals Hotelling chart or CUSUM or EWMA charts for more variables (Noskievičová, 2008).

4. RESEARCH METHOD CONDUCTED

As the title of the paper explains, the primary method used for achieving the goals is the case studies research. The purpose of these studies is to illustrate how time series charts are found to be sensitive in detecting small shifts and we will utilize the fact that these control charts can be used in certain situations where the data are autocorrelated. The methodology combined quantitative data analysis with financial process simulation, which benefits the reliability of the conclusions.

5. RESULTS OF THE CASE STUDIES

5.1 Case Study No 1 – A Use of the CUSUM Control Chart

- $\mu_0 = 10, n = 1, \sigma = 1,0$
- We would like to detect the shift $1.0\sigma = 1.0(1.0) = 1.0$ ($d = 1,0$)
- A process mean which is out of control: $\mu_1 = 10 + 1 = 11$
- $K = d/2 = 1/2$ a $H = 5\sigma = 5$ (recommended)
- Equations for the statistics C_i^+ and C_i^- are then:

$$C_i^+ = \max\left[0, x_i - 10.5 + C_{i-1}^+\right] \tag{16}$$

$$C_i^- = \max\left[0, 10.5 - x_i + C_{i-1}^-\right]$$

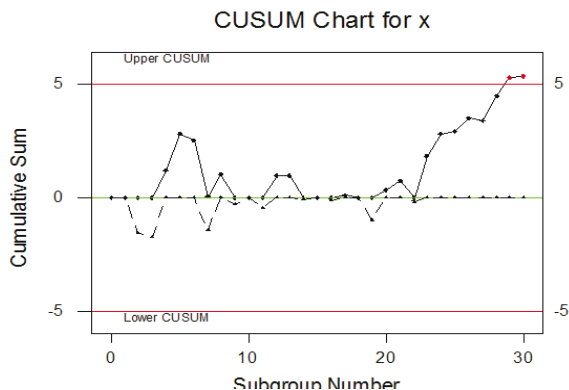


Fig. 3 – The Control Chart CUSUM. Source: QC Expert 2.5C \bar{x}

The CUSUM chart shows that the process is out of control (see Figure 3). Following steps are to find the main cause (-s), accept precaution (-s) and to start a new CUSUM again. If the process was adjusted it could be useful to estimate mean of the process caused by the shift.

5.2 Case Study No 2 – A Use of Control Chart EWMA

We have following financial data of a certain company (in millions of CZK). The starting value of x is 5.00. Now, we construct a control chart for data from the Tab. 1.

Tab. 1 – Entering Data for Control Chart EWMA. Source: Own Processing

K	x_k	k	x_k	k	x_k	k	x_k	k	x_k
1	5.01	11	4.85	21	4.80	31	4.77	41	4.65
2	5.11	12	4.99	22	4.84	32	4.60	42	4.67
3	5.04	13	5.05	23	4.78	33	4.51	43	4.50
4	5.12	14	5.38	24	4.82	34	4.58	44	4.50
5	4.94	15	4.97	25	4.88	35	4.67	45	4.44
6	5.01	16	4.84	26	4.78	36	4.51	46	4.44
7	5.11	17	4.77	27	4.80	37	4.57	47	4.53
8	5.18	18	4.78	28	4.81	38	4.56	48	4.53
9	5.04	19	4.82	29	4.88	39	4.57	49	4.57
10	5.18	20	4.78	30	4.75	40	4.60	50	4.34

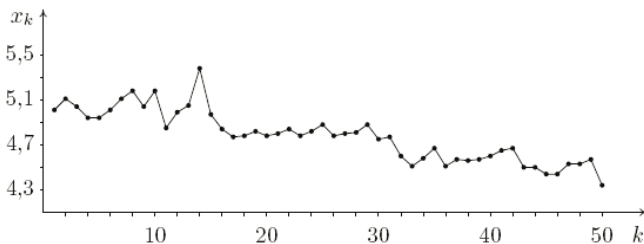


Fig. 4 – Lineplot. Source: QC Expert 2.5C \bar{x}

It is evident from the Figure 4 that the values of the process are declining – therefore there is not a constant process mean. Further, we find out if there is an autocorrelation of the 1st degree in the process. We construct a correlation chart between the x_k and x_{k+1} values for $k = 1, 2, 3, \dots, 49$ (see Figure 5).

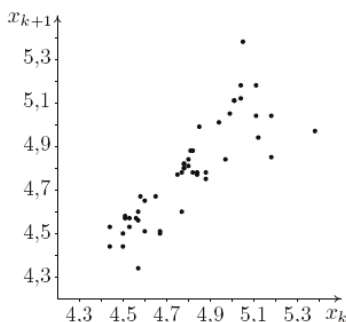


Fig. 5 – Correlation chart between the x_k and x_{k+1} values (scatterplot). Source: QC Expert 2.5C \tilde{x}

We can see from the Figure 5 now that the scatterplot is ellipse-shaped and that the ellipse main axis forms an acute angle with the x-axis. We can conclude from this exploratory data analysis that there is a significant first degree positive autocorrelation between the variables x_k and x_{k+1} . We calculate exact coefficient of autocorrelation between the variables x_k and x_{k+1} . It is equal to 0.850. Thus, the strong (statistically significant) positive autocorrelation exists between the variables x_k and x_{k+1} . Therefore we can construct the EWMA dynamic control chart. Further, we will compute predicted values for the empirically chosen values of the parameter λ from the interval $(0;1)$. A starting value was set to the number $\mu_0 = 5.00$. We determine the value of function (the sum of error squares) for chosen parameter λ at the same time. The calculation of the parameter $\lambda = 0.48$ is shown in Table 2. This value was found as optimal. Function $S(\lambda)$ is evidently parabolic with one extreme (minimum).

Tab. 2 – Values of Function $S(\lambda)$. Source: Own Processing

λ	$S(\lambda)$	λ	$S(\lambda)$
0.4	0.663138	0.47	0.658893
0.5	0.659108	0.48	0.658852
0.6	0.665795	0.49	0.658926

The Table 2 shows computing of the optimal λ parameter value. The values of λ can be found in the first and the third columns and the values of $S(\lambda)$ function in the second and the fourth columns. The values of $S(\lambda)$ function for 0.4, 0.5, and 0.6 are given in the first two columns of the Table 2. We can see that the function minimum is between values 0.4 and 0.5. Therefore we try values of λ equal to 0.47, 0.48 and 0.49. According to the previous calculations, the function minimum $S(\lambda)$ is equal to 0.48.

The final EWMA control chart is displayed on the Figure 6 where lower and upper control limits are plotted with bold full black lines, center line is plotted with line of dashes. Input values of quality feature are represented by the bold dots.



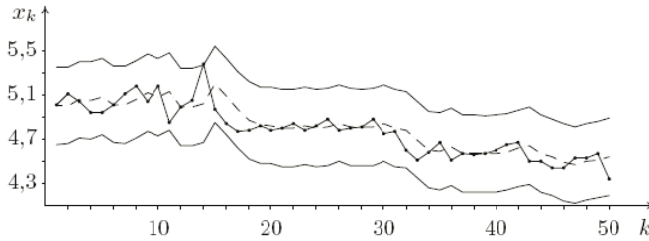


Fig. 6 – Control Chart EWMA. Source: QC Expert 2.5Cz

We can conclude from the analysis of the Figure 6 that the process is under control except of the 14th value. Other values are between the control limits. The 14th value crossing can be caused by the process nonstability or by the measurement inaccuracy.

5.3 Case Study No 3 – A Use of Control Charts ARIMA CUSUM and EWMA

We find a suitable model for the time series of annual gross domestic product of the United Kingdom which is available in the form of base indices from 1960 to 1997 (1995 = 100). We determine the control chart sensitive to a change of the mean based on this model. Financial data are in the following table.

Tab. 3 – Financial Data. Source: Own Processing

t	x_t	t	x_t	t	x_t	t	x_t
1	45	11	59	21	72	31	94
2	46	12	61	22	71	32	92
3	47	13	63	23	72	33	91
4	49	14	67	24	75	34	93
5	51	15	66	25	77	35	97
6	53	16	66	26	80	36	100
7	54	17	68	27	83	37	102
8	55	18	69	28	87	38	106
9	57	19	72	29	91		
10	58	20	74	30	93		

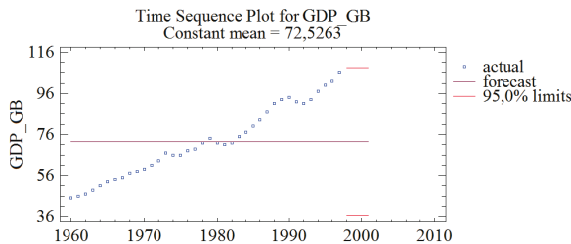


Fig. 7 – Time Series of Automated Production Process. Source: Own Processing

The time series is shown on the previous figure 7. It is obvious that the time series is nonstationary, which is also confirmed by the shape of the ACF and PACF (Fig. 8) and periodogram (Fig. 9)

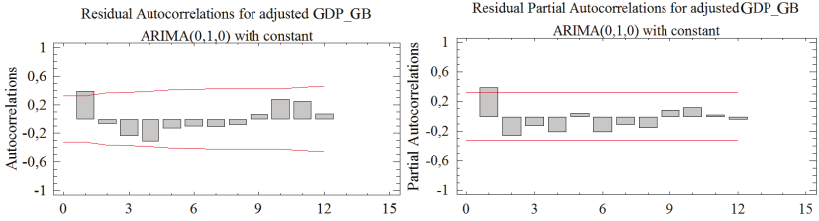


Fig. 8 – ACF and PACF of Original Time Series. Source: Own Processing

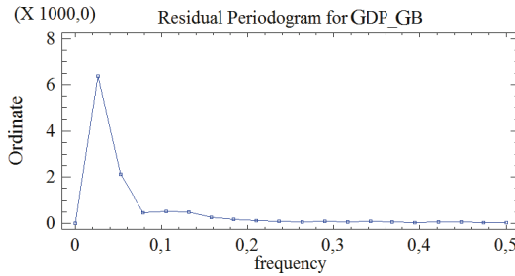


Fig. 9 – Periodogram of Original Time Series. Source: Own Processing

Because the values of ACF are declining slowly, the first values by ACF and PACF are very close to one and the periodogram has a significant peak at zero frequency, we can assume that the time series is of a Type I (1) and therefore must be stationarized by the first differences. This conclusion is confirmed by the graph of residuals and residual periodogram (Figure 10), as well as the residual ACF and PACF of the model (Fig. 11), whose values lie within the tolerance limits.

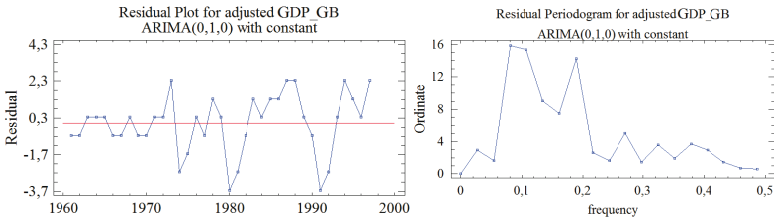


Fig. 10 – Residuals and Residual Periodogram of Model ARIMA(0,1,0). Source: Own Processing

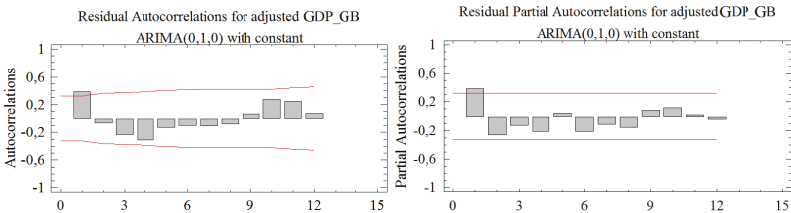


Fig. 11 – Residual ACF and PACF of Model ARIMA(0,1,0). Source: Own Processing

Tab. 4 – Output Table of Model ARIMA(0,1,0)_c. Source: Own Processing

Nonseasonal differencing of order: 1			Forecast model selected: ARIMA(0,1,0) with constant				
Number of forecasts generated: 4			Number of periods withheld for validation: 0				
	<i>Estimation</i>	<i>Validation</i>	ARIMA Model Summary				
<i>Statistic</i>	<i>Period</i>	<i>Period</i>	<i>Parameter</i>	<i>Estimate</i>	<i>Std. Error</i>	<i>t</i>	<i>P-value</i>
RMSE	1,60236		Mean	1,64865	0,263427	6,25847	0,000000
MAE	1,22863		Constant	1,64865			
MAPE	1,62859		Backforecasting: yes				
ME	3,07262E-15		Estimated white noise variance = 2,56757 with 36 degrees of freedom				
MPE	-0,106827		Estimated white noise standard deviation = 1,60236				
Tests for Randomness of residuals							
Box-Pierce Test							
Test based on first 12 autocorrelations							
Large sample test statistic = 18,8926							
P-value = 0,0911537							

Residual ACF and PACF have the first values statistically significantly different from zero. Therefore cannot be clearly determined whether the model should be extended to include the part of the AR(1) or MA(1). First, we estimate ARIMA(1,1,0)_c model. The results are presented in Table 5.

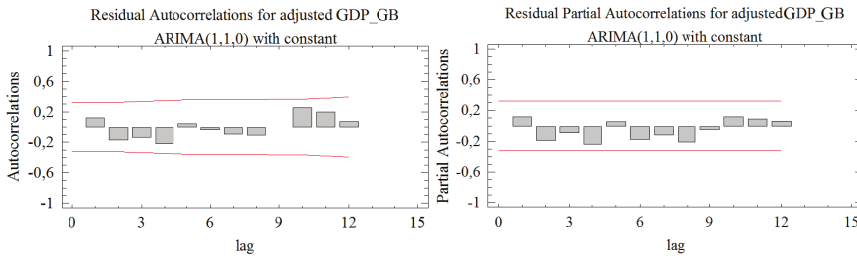


Fig. 12 – Residual ACF and PACF of Model ARIMA(1,1,0)_c. Source: Own Processing

Although residual ACF and PACF indicate that non-systematic component is a type of white noise, as well as the value of “P-value” of the Box-Pierce test is relatively high (0,597248), we will also estimate the model ARIMA(0,1,1)_c. The Table 5 contains its results.

Tab. 5 – Output Table of Model ARIMA (1,1,0)_c Source: Own Processing

Nonseasonal differencing of order: 1			Forecast model selected: ARIMA(1,1,0) with constant				
Number of forecasts generated: 4			Number of periods withheld for validation: 0				
	<i>Estimation</i>	<i>Validation</i>	ARIMA Model Summary				
<i>Statistic</i>	<i>Period</i>	<i>Period</i>	<i>Parameter</i>	<i>Estimate</i>	<i>Std. Error</i>	<i>t</i>	<i>P-value</i>
RMSE	1,48723		AR(1)	0,407307	0,159438	2,55464	0,015137
MAE	1,09575		Mean	1,66965	0,394347	4,23396	0,000158
MAPE	1,46969		Constant	0,989589			
ME	0,0162096		Backforecasting: yes				
MPE	-0,0476283		Estimated white noise variance = 2,21333 with 35 degrees of freedom				
Number of iterations: 1							
Tests for Randomness of residuals							
Box-Pierce Test							
Test based on first 12 autocorrelations							
Large sample test statistic = 9,26715							
P-value = 0,597248							

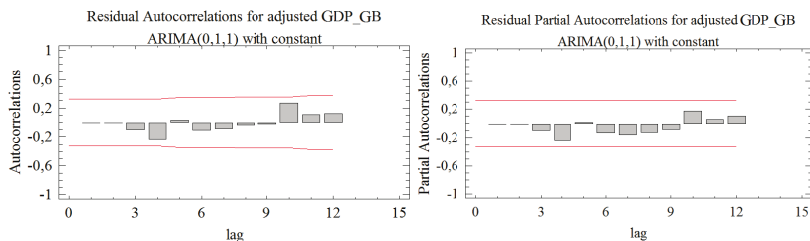


Fig. 13 – Residual ACF and PACF of Model ARIMA(0,1,1)_c. Source: Own Processing

Residual ACF and PACF of the model also indicate that the type of non-systematic component is white noise, the value of “P-value” of the Box-Pierce test (0,794106) is even higher than before. For completeness, we estimate another model containing two components, i.e. the AR(1) and MA(1).

Tab. 6 – Output Table of Model ARIMA(0,1,1)_c Source: Own Processing

Nonseasonal differencing of order: 1		Forecast model selected: ARIMA(0,1,1) with constant		Number of forecasts generated: 4		Number of periods withheld for validation: 0	
		ARIMA Model Summary					
<i>Statistic</i>	<i>Estimation</i>	<i>Validation</i>	<i>Parameter</i>	<i>Estimate</i>	<i>Std. Error</i>	<i>t</i>	<i>P-value</i>
	<i>Period</i>	<i>Period</i>	MA(1)	-0,569874	0,147994	-3,85066	0,000480
RMSE	1,43305		Mean	1,67905	0,362275	4,63471	0,000048
MAE	0,9854		Constant	1,67905			
MAPE	1,32419		Backforecasting: yes				
ME	0,00350342		Estimated white noise variance = 2,054 with 35 degrees of freedom				
MPE	-0,0689866		Estimated white noise standard deviation = 1,43318				
				Number of iterations: 5			
Tests for Randomness of residuals							
Box-Pierce Test							
Test based on first 12 autocorrelations							
Large sample test statistic = 7,06119							
P-value = 0,794106							

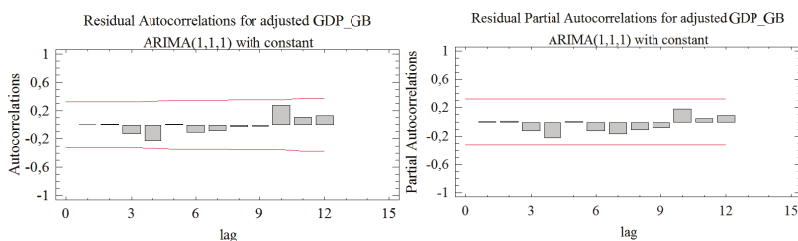


Fig. 14 – Residual ACF and PACF of Model ARIMA(1,1,1)_c. Source: Own Processing

Although residual ACF and PACF in Figure 14 indicate that the residuals are non-systematic and “P-value” of the Box-Pierce test is relatively high (0,701229), it is clear from the t-tests of the model parameters that the AR(1) part is not a part of the model.

However, not only this conclusion is in favor of the model ARIMA(0,1,1)_c. This model leads to a higher value of “P-value” of the Box-Pierce test and to lower values of MSE, MAE, MAPE and MPE in comparison with the model ARIMA (1,1,0)_c.

Tab. 7 – Output Table of Model ARIMA(1,1,1)c Source: Own Processing

Nonseasonal differencing of order: 1			ARIMA Model Summary				
Forecast model selected: ARIMA(1,1,1) with constant			<i>Parameter</i>	<i>Estimate</i>	<i>Std. Error</i>	<i>t</i>	<i>P-value</i>
Number of forecasts generated: 4			AR(1)	-0,0807671	0,312105	-0,258782	0,797354
Number of periods withheld for validation: 0			MA(1)	-0,635226	0,246258	-2,57952	0,014390
<i>Statistic</i>	<i>Estimation Period</i>	<i>Validation Period</i>	Mean	1,6766	0,361225	4,64143	0,000050
RMSE	1,4525		Constant	1,81201			
MAE	0,98844		Backforecasting: yes				
MAPE	1,32751		Estimated white noise variance = 2,10992 with 34 degrees of freedom				
ME	0,00204376		Estimated white noise standard deviation = 1,45256				
MPE	-0,0731281		Number of iterations: 6				
Tests for Randomness of residuals			Box-Pierce Test				
Test based on first 12 autocorrelations			Large sample test statistic = 7,25441				
P-value = 0,701229							

After comparing the values of “P-value” of the parameters $c = \mu$ (0,000048) and θ_1 (0,000480) with a significance level α (0,05) by a t-test of the Table 6, it is evident that both parameters are non-zero and the estimated model has the form

$$(1-B)y_t = 1.67905 + (1 - (0.569874)B)a_t,$$

can be expressed as:

$$y_t = 1.67905 + y_{t-1} + a_t + 0.569874 a_{t-1}.$$

Forecasts are shown in Figure 15 and a time series plot with smoothed values and predictions is shown in Figure 16.

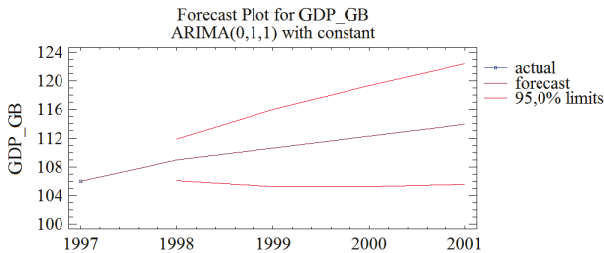


Fig. 15 – Forecasts. Source: Own Processing

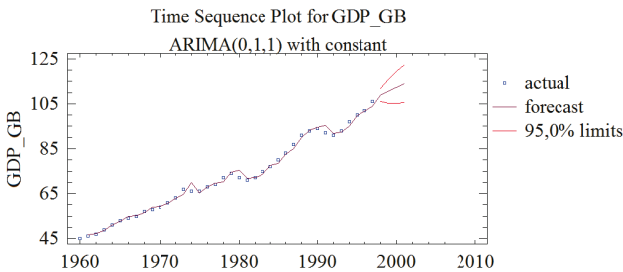


Fig. 16 – Time Series with Forecasts. Source: Own Processing

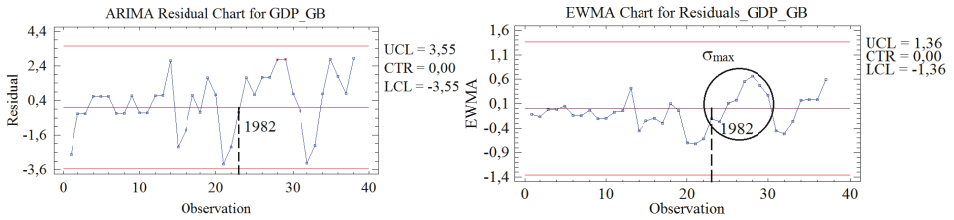


Fig. 17 – Control Chart for Individual Values Applied on Residuals of Chosen ARIMA Model (left) and EWMA Control Chart for Residuals of ARIMA Model with $\lambda = 0.2$ (right). Source: Own Processing

From those control charts (Figure 17) can be seen that the process can be regarded as if it is statistically stable (no point is outside the control limits), so these limits can be used to verify the statistical stability of the process in the next period.

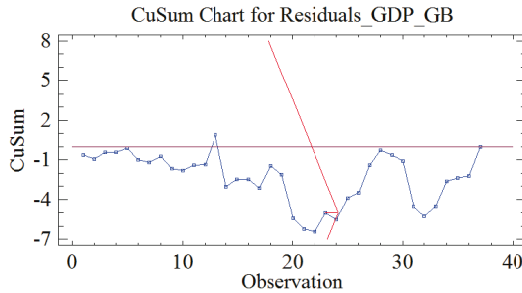


Fig. 18 – CUSUM Chart for Residuals. Source: Own Processing

The CUSUM control chart shows that the V-mask identifies the beginning of increasing heteroscedasticity in point 23, which corresponds to a year 1982 (Figure 18). Previous control charts show the same changing point in the volatility of the time series and due to increasing heteroscedasticity indicate a growing trend of time series.

6. DISCUSSION

Most traditional control charting procedures are grounded on the assumption that the process observations being monitored are independent and identically distributed. With the advent of high-speed data collection schemes, the assumption of independence is usually violated. That is, autocorrelation among measurements becomes an inherent characteristic of a stable process. This autocorrelation causes significant deterioration in control charting performance. To address this problem, several approaches for handling autocorrelated processes have been proposed. The most popular procedure utilizes either a Shewhart, CUSUM or EWMA chart of the residuals of the appropriately fitted ARMA model. However, procedures of this type possess poor sensitivity especially when dealing with positively autocorrelated processes. As an alternative, we have explored the application of the statistics used in a time series procedure for detecting outliers and level shifts in process monitoring. The study focused on the detection of level shifts of autocorrelated processes with particular emphasis on the important AR(1) model. The

results presented showed that time series charts are found to be sensitive in detecting small shifts and we utilize the fact that these control charts can be used in certain situations where the data are autocorrelated.

7. CONCLUSION

This paper dealt with the regulation charts applications in financial data. This kind of data is very sensitive to mean shifting and strong autocorrelation appears very often. Therefore we put a focus on dynamic regulation charts CUSUM, EWMA and ARIMA models. We highlighted the versatility of control charts not only in manufacturing but also in managing the financial stability of cash flows in this paper. A refined identification of the type of intervention affecting the process will allow users to effectively track the source of an out-of-control situation which is an important step in eliminating the special causes of variation. It is also important to note that the proposed procedure can also be applied when dealing with a more general autoregressive integrated moving average model. Autocorrelated process observations mainly arise under automated data collection schemes. Such collection schemes are typically controlled by software which can be upgraded to handle SPC functions. Under such an integrated scheme the usefulness of the proposed procedure will be optimized. Based on information from chapter 5, we would recommend a properly designed time series control charts as a control charts for individual measurements in a wide range of applications. These are almost a perfectly nonparametric (distribution-free) procedures.

Acknowledgement

This paper has been supported by IGA TBU No IGA/73/FaME/10/, Development in usage of mathematical and statistical methods in quality management.

References

1. Atienza, O. O., Tang, L. C. & Ang, B. W. (1997). ARL properties of a sample autocorrelation chart. *Computers and Industrial Engineering*, 33 (3-4), 733-736. [http://dx.doi.org/10.1016/S0360-8352\(97\)00234-9](http://dx.doi.org/10.1016/S0360-8352(97)00234-9)
2. Alwan, L. C. (1992). Effects of autocorrelation on control chart performance. *Communications in Statistics: Theory and Methods*, 21 (4), 1025-1049. <http://dx.doi.org/10.1080/03610929208830829>
3. Apley, D. W. & Tsung, F. (2002). The autoregressive T2 chart for monitoring univariate autocorrelated processes. *Journal of Quality Technology*, 34 (1), 80-96.
4. Dyer, J. N., Conerly, D. M. & Adams, B. M. (2003). A simulation study and evaluation of multivariate forecast based control charts applied to ARMA processes. *Journal of Statistical Computation and Simulation*, 73 (10), 709-724. <http://dx.doi.org/10.1080/0094965031000062168>
5. Gervini, D. (2003). A robust and efficient adaptive reweighted estimator of multivariate location and scatter. *Journal of Multivariate Analysis*, 84 (1), 116-144. [http://dx.doi.org/10.1016/S0047-259X\(02\)00018-0](http://dx.doi.org/10.1016/S0047-259X(02)00018-0)
6. Grubbs, F. E. (1969). Procedures for Detecting Outlying Observations in Samples. *Technometrics*, 11 (1), 1-21. <http://dx.doi.org/10.1080/00401706.1969.10490657>

7. Harris, T. J. & Ross, W. H. (1991). Statistical process control procedures for correlated observations. *Canadian Journal of Chemical Engineering*, 69 (1), 48-57. <http://dx.doi.org/10.1002/cjce.5450690106>
8. Hušek, R. (2007). *Ekonomická analýza*. Praha: Vysoká škola ekonomická v Praze, Nakladatelství Oeconomica.
9. Chambers, D. S. & Wheeler, D. J. (1992). *Understanding Statistical Process Control*. USA: SPC Press, Inc.
10. Chandra, M. J. (2001). *Statistical Quality Control*. United States of America: CRC Press, LLC. <http://dx.doi.org/10.1201/9781420038675>
11. Krieger, C. A., Champ, C. W. & Alwan, L. C. (1992). *Monitoring an autocorrelated process*. Presented at the Pittsburgh Conference on Modeling and Simulation.
12. Lu, C. W. & Reynolds, M. R. (1999). Control charts for monitoring the mean and variance of autocorrelated processes. *Journal of Quality Technology*, 31, 259-274.
13. Lu, C. W. & Reynolds, M. R. (1999). EWMA control charts for monitoring the mean of autocorrelated processes. *Journal of Quality Technology*, 31 (2), 166-188.
14. Meloun, M. & Militký, J. (2006). *Kompendum statistického zpracování dat*. Praha: Academia, nakladatelství Akademie věd České republiky.
15. Montgomery, D. C., & Friedman, D. J. (1989). *Statistical process control in computerintegrated manufacturing environment*. In J. B. Keats & N.F. Hubele (eds). *Statistical Process Control in Automated Manufacturing*. New York: Marcel Dekker.
16. Montgomery, D. C. & Mastrangelo, C. M. (1991). Some statistical process control methods for autocorrelated data (with discussion). *Journal of Quality Technology*, 23 (3), 179-204.
17. Noskiewiczová, D. (2008). Vybrané metody statistické regulace procesu pro autokorelovaná data. *AUTOMA*, 10, 40-43.
18. Runger, G. C. & Willemain, T. R. (1996). Batch-means control charts for autocorrelated data. *IIE Transactions*, 28 (6), 483-487. <http://dx.doi.org/10.1080/07408179608966295>
19. Sun, J. & Xu, L. (2004). Batch average control chart. *ASQ Annual Quality Congress Proceedings*, 58.
20. Yourstone, S. A. & Montgomery, D. C. (1991). Detection of process upsets – sample autocorrelation control chart and group autocorrelation control chart applications. *Quality and Reliability Engineering International*, 7 (3), 133-140. <http://dx.doi.org/10.1002/qre.4680070304>

Contact information

Ing. Bc. Martin Kovářik, Ph.D.

Assoc. Prof. Ing. Petr Klímeček, Ph.D.

Tomas Bata University in Zlín, Faculty of Management and Economics

Department of Statistics and Quantitative Methods

Mostní 5139, 760 01 Zlín

Tel: +420 576 032 815

E-mail: kovarik.fame@seznam.cz

E-mail: klimecek@fame.utb.cz

JEL Classification: C39, C49, C59
