Physicochem. Probl. Miner. Process. 48(2), 2012, 495−512

www.minproc.pwr.wroc.pl/journal/

Physicochemical Problems of Mineral Processing

ISSN 1643-1049 (print)
ISSN 2084-4735 (online)

Received April 3, 2012; reviewed; accepted May 1, 2012

EFFECT OF AUTOCORRELATION ON THE PROCESS CONTROL CHARTS IN MONITORING OF A COAL WASHING PLANT

Adem TASDEMIR

Department of Mining Engineering, Division of Mineral Processing, Eskişehir Osmangazi University, 26480, Eskişehir, Turkey, atasdem@ogu.edu.tr, Tel: +90 222 239 37 50 Ext: 3438

Abstract. Traditional statistical process control charts assume that generated process data are normally and independently distributed, i.e. uncorrelated. This research presents the effect of autocorrelation on process control charts to monitor the two quality characteristics of fine coals produced in a coal washing plant for power plant, namely moisture content and ash content. Individual (X) and moving range charts (MR) were constructed to monitor 10 months data. It was determined that even though both data values obey the normal distribution, there is a moderate autocorrelation between their observations. For simulating the autocorrelated data, ARIMA time-series models were used. It was found that X/MR charts showed many false alarms due to the autocorrelation. The ARIMA (1, 0, 1) for moisture content and ARIMA (0, 1, 2) for ash content were determined to be the best models to remove autocorrelation. Compared to large number of false alarms on conventional X/MR charts and on charts applying the Western Electric rules, which assume the data independence, there were much less unusual points on the X/MR charts of residuals (Special Cause Charts). Usage of residual based control charts is suggested when the data are autocorrelated.

keywords: coal, autocorrelation, statistical control charts, special cause charts, ARIMA models

1. Introduction

Statistical process control (SPC) has been proposed firstly by Dr. Shewhart in 1931 and then, many SPC charts have been developed and improved to use for different process data. In its basics form, a control chart compares process observations with a pair of control limits (Karaoğlan and Bayhan, 2011). One of the essentials of producing high quality low cost product is to adopt and apply the SPC correctly. SPC is a tool for achieving and improving quality standards. The most important and sophisticated tool of SPC is control charts.

SPC is widely used to monitor, control and improve quality in many industrial processes (MacCharty and Wasusri, 2002). The SPC has been also used in mining and mineral processing field to monitor the grade variations and to ensure the product specifications. In a review study, MacCharty and Wasusri (2002) summarize the usage of non-standard applications of statistical process control charts. Some application

http://dx.doi.org/10.5277/ppmp120215

examples of SPC can be given from the literature for mining/mineral processing industry (Ankara and Bilir, 1995; İpek et al., 1999; Bhattacherjee and Samanta, 2002; Bayat and Arslan, 2004; Aykul et al., 2005; Vapur et al., 2005; Elevli and Behdioğlu, 2006; Ankara and Yerel, 2008; Vapur, 2009; Elevli, 2009; Aykul et al., 2010; Yerel and Ankara, 2011). However, the information related to the application of SPC in mineral industry is still limited when compared to other industrial areas.

Traditional SPC techniques are based on the assumption that generated data generated are normally and independently distributed. However, the independency assumption is not realistic in practice (Demirkol, 2008). Therefore, the assumption of normality and indepence of data must be satisfied in order to apply SPC for a process (Bhattacherjee and Samanta, 2002).

In the continuous industries processes most data are autocorrelated (Le, 1998). Since observations for many processes exhibit autocorrelation that may be the result of dynamics that are inherent to the process, the autocorrelation is more likely to be observed in processes when observations are closely spaced in time (Lu and Reynolds, 2001). Such autocorrelations can occur for a number of reasons. One reason is because of current automated measurement and recording technology, subgroup samples may be taken with a high frequency, with consecutive samples being similar in nature and hence statistically correlated. Other samples of correlated subgroups occur when items made by a worker exhibit similar characteristics due to the way the machine is handled, the process shows seasonal patterns due to materials or weather, or the alertness of a worker changes over time (Le, 1998).

The effect of autocorrelation on the performance of process control charts have been shown by many researhers (Alwan and Roberts, 1998; Lu and Reynolds, 1999; Montgomery and Mastrangeko, 1991; Atienza et al., 1997; Reynolds and Lu, 1997; Zhang, 1997). Results of these studies have shown that traditional SPC methods are strongly affected by data autocorrelation. Under autocorrelation conditions, the traditional control charts will become ineffective and many false alarms may occur on the charts. Consequently, there has been considerable research in recent years on designing SPC procedures suitable for autocorrelated processes.

In the case of autocorrelation, a very high false alarm rate will cause process personnel to waste effort in unproductive searches for special causes. This can lead to a loss of confidence in the control chart, and even to process monitoring being discontinued. Thus, autocorrelation should not be ignored when designing control charts, because failure to properly account for autocorrelation can greatly reduce or eliminate the effectiveness of control charts (Lu and Reynolds, 2001).

For these reasons, some modifications for traditional control charts are necessary. As an alternative to traditional control charts which plot the original observations, a number of papers have suggested fitting a time series model (Alwan and Roberts, 1988; Reynolds and Lu, 1997; Lu and Reynolds, 1999; Lu and Reynolds, 2001). The residuals obtained from time series model are then plotted on standard control charts.

To apply these methods, a time series model of the process is required (Apley and Lee, 2003).

Alwan and Roberts (1988) introduced the first residual control chart, namely special cause chart (SCC). A residual Shewhart chart was developed by them and it is called as the special control chart (SCC). The SCC is an individual control chart (individual control chart is constructed by plotting single observations and setting control limits on these observations) applied to the residuals. Therefore, the SCC is also known as the X residual chart in the literature. In residual charts, an appropriate time series model is fitted to the autocorrelated observations and the residuals are plotted in a control chart (Demirkol, 2008).

Even though the assumptions required for the applicability of SPC methods are known to be essential, their existence may questionable in mining industry (Bhattacherjee and Samanta, 2001). When the literature is reviewed, the impact of autocorrelation on SPC have been considered in mineral applications by only few studies (Samanta and Bhattacherjee, 2001; Bhattacherjee and Samanta, 2002; Samanta, 2002; Elevli et al., 2009). Indeeed, the observations in these studies are assumed to normally distributed and independent that is uncorrelated.

Turkey has 8.3 billion tons of lignite coal reserves (Arslan, 2006). Since 80% of Turkish lignite coal reserves has a heating value of below 2500 kcal/kg, large amount of them are used in power plants. Moisture and ash contents of coals are the two important quality characteristics that determine properties of coals burning in power plants. Substantial amount of low quality coals which have high ash and moisture contents are transformed into energy by burning them in power plants (Tekir et al., 2009). In a power plats that burns lignites, moisture and ash contents are the key quality indicators monitored. Therefore, lowering the ash contents of lignites to the desired levels is an essential entity to reduce problems and costs associated with the ash. Coals having high ash and moisture contents cause serious problems such as incrasing the costs of operation, transportation and burning of coal together with disposing the ash to dams in power plants (Tekir et al., 2009). For these reasons, coal fines produced in a coal washing plant should have desired moisture and ash content limits in order to be used as feeding material for a power plant. To control the coal qualities produced in washing plants, it is necessary to monitor the coal characteristics by applying an effective method such as statistical process control.

The aim of this paper is to investigate individual X control charts based on the original observations or on residuals obtained by the time series of ARIMA models for moisture and ash contents of fine coal products that are produced for power plant as a feeding material in Tunçbilek washing plant. Autocorrelated observations were characterized by ARIMA models and special control charts (SCC) were utilized to monitor the residuals, based on the ARIMA model forecast values. The performance of the standard Shewhart chart, which ignores subgroup correlations for the two coal quality variables, is compared with the X charts of residuals.

2. Statistical process control charts

Statistical Process Control (SPC) charts are graphical presentations of the sample quality for one parameter measured in control samples. In a basic the SPC chart, three horizontal lines, the center line and two control limits, are plotted on a control chart, which show the mean value and three standard deviations distance from either side of the mean value, respectively (an example is given in Fig. 1). If the plotted observatios fall outside the limits, the process is consired to be out of control. The process needs to be stopped and inspected for causes when the out of control points are detected (Carson and Yeh, 2008).

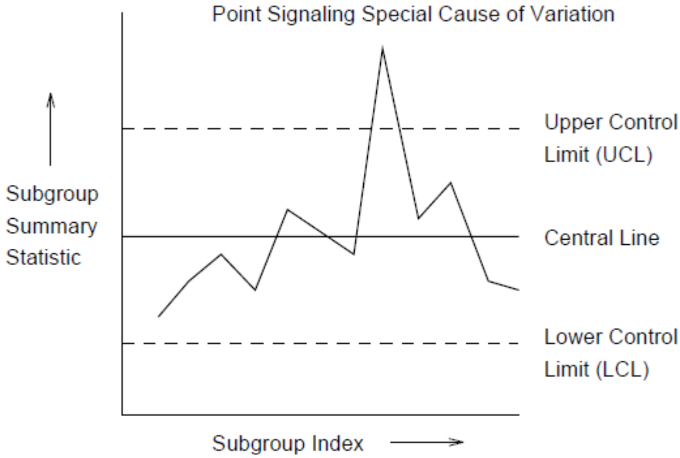

Fig. 1. A traditional Shewhart control chart

2.1. X/MR charts for individual observations

In many situations, the sample size used for process control is n = 1; that is, the sample consists of an individual unit. In such situation, the individuals control chart (also called as X chart or I chart) is used. Mean of the individual values for m individual observations gives us the center line of the X control chart as:

$$\overline{X} = \frac{1}{m}\sum_{i=1}^{m} X_i = \frac{X_1 + X_2 + X_3 + \ldots + X_m}{m}. \tag{1}$$

The control chart for individuals uses moving range of two successive observations to estimate the process variability and it is also known as MR chart (Montgomery and Runger, 2011). The moving range is defined as $MR_i = |X_i\text{-}X_{i\text{-}1}|$ and for m observations the average moving range ($\overline{MR}$) is:

$$\overline{MR} = \frac{1}{m-1}\sum_{i=2}^{m} \left|X_i - X_{i-2}\right|. \tag{2}$$

An estimate of standard deviation (σ) is:

$$\sigma = \frac{\overline{MR}}{d_2} = \frac{\overline{MR}}{1.128}. \tag{3}$$

Here, d_2 is a constant depending on the sample size, n, (Montgomery and Runger, 2011). It is 1.128 for $n = 2$.

The center line (CL_X) and upper control limit (UCL_X) and lower control limit (LCL_X) for a control chart for individuals is (Montgomery and Runger, 2011):

$$UCL_X = \overline{X} + 3\frac{\overline{MR}}{d_2} = \overline{X} + 3\frac{\overline{MR}}{1.128} = \overline{X} + 2.66\overline{MR}, \tag{4}$$

$$CL_X = \overline{X}, \tag{5}$$

$$LCL_X = \overline{X} - 3\frac{\overline{MR}}{d_2} = \overline{X} - 3\frac{\overline{MR}}{1.128} = \overline{X} - 2.66\overline{MR}. \tag{6}$$

Since each moving range is the range between two consecutive observations. It should be noted that there are only m-1 moving ranges. The parameters for a control chart of moving range are defined as follows:

$$UCL_{MR} = D_4\overline{MR} = 3.267\overline{MR}, \tag{7}$$

$$CL_{MR} = \overline{MR}, \tag{8}$$

$$LCL_{MR} = D_3\overline{MR} = 0. \tag{9}$$

D_3 and D_4 are constants depending on the sample size, n. It is possible to establish a control chart on the MR using D_3 and D_4 for $n = 2$ (Montgomery and Runger, 2011). The LCL_{MR} for this moving range chart is always zero because $D_3 = 0$ for $n = 2$. The D_4 equals to 3.267 for $n = 2$ while determining the UCL_{MR} .

For a simple control chart in Fig. 1 control limits were plotted as warning lines from the action limits of 3σ. In this figure, only measurements that are plotted further than 3σ from the central line indicate an out-of-control process. However, additional warning limits in Shewhart charts are commonly used in practice, since they help the analyst to recognize possible changes in the process before it shifts out of 3σ (Carson and Yeh, 2008). For example, when two consecutive out of 2σ samples are detected, this can indicate some changes in the process. Therefore, it is advisable to perform some rechecks of the process at this point.

It is known that the Shewhart control chart is insensitive to small shifts in the process parameter. To enhance the ability for the chart to detect small shifts more quickly, one way is to add sensitive rules in the chart (Chang and Wu, 2011). Several frequently used rules suggested by the Western Electric (1956) are (Montgomery and Runger, 2011):

1. One or more points fall outside the 3σ control limits,
2. Two of three consecutive points fall outside the 2σ warning limits,
3. Four of five consecutive points fall beyond the 1σ limits,
4. Eight points in a row fall one side of the center line.

The above rules are always referred to runs rules. In the sequel, a control chart that uses several rules simultaneously is referred to a compound control chart. It should be noted that rule 1 in included in all compound control charts. Also note that these rules can only be applied to one side of the center line. There are more rules for detecting small shifts and these can be added to control.

In this study, the four rules listed above were used when considering additional warning limits by Western electric rules. These plots searches and identifies any unusual patterns in the data. This is often helpful in detecting processes which are slowly drifting away from target, even though no points may fall outside the control limits.

3. Time series models

The methodology of ARIMA estimation and model selection is a classical topic covered in most textbooks on time series analysis such as Montgomery et al. (2008). Therefore, we will not duplicate the descriptions of already well documented methodologies. We will give only practical information in this context to the models.

3.1. Autoregressive Process (AR)

Most time series consist of elements that are serially dependent in the sense that one can estimate a coefficient or a set of coefficients that describe consecutive elements of the series from specific, time-lagged (previous) elements. This can be summarized in the following Eq. 10:

$$X_t = \delta + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots + \phi_p X_{t-p} + \varepsilon_t , \tag{10}$$

where X_t is response (dependent) variable at time t; $X_{t\text{-}1}$, $X_{t\text{-}2}$, $X_{t\text{-}p}$ response variable at time lags t-1, t-2,…t-p, respectively; ϕ_1, ϕ_2, ϕ_p, denote estimated coefficients, ε_t is an error term at time t, and $\delta = \left(1 - \sum_{i=1}^{p} \phi_i\right)\mu$ with μ denotes the process mean.

It should be noted that an autoregressive process will only be stable if the parameters are within a certain range for example, if there is only one autoregressive parameter, then it must fall within the interval of $-1<\phi_1<+1$. Otherwise, past effects would accumulate and the values of successive X_t's would move towards infinity, that is, the series would not be stationary. If there is more than one autoregressive parameter, similar (general) restrictions on the parameter values can be defined (Montgomery et al., 2008).

3.2. Moving Average process (MA)

Independent from the autoregressive process, each element in the series can also be affected by the past error (or random shock) that cannot be accounted for by the autoregressive component, that is:

$$X_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q}\,, \tag{11}$$

where X_t is the time series, μ is the mean of the series, ε_{t-I} denotes random noise term with mean 0 and variance σ_ε^2 at time *t*. θ_1, θ_2, θ_q, are the moving average model parameters. The value of *q* is called the order of the MA model.

3.3 Autoregressive Moving Average model (ARIMA)

The ARIMA model which combines the autoregressive and moving average parameters and includes differencing in the model was introduced by Box and Jenkins (1976) (Montgomery et al., 2008). This equation was used to forecast one-step ahead disturbances, according to data characteristics of stationary or non-stationary as shown in Eq. 12.

$$\Delta_d X_t = \mu + \phi_1 \Delta_d X_{t-1} + \phi_2 \Delta_d X_{t-2} + \ldots + \phi_p \Delta_d X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \ldots - \theta_q \varepsilon_{t-q}\,, \tag{12}$$

Three types of parameters in the model are: the autoregressive parameters (*p*), the number of differencing passes (*d)*, and moving average parameters (*q*). The Box and Jenkins models are summarized as the ARIMA (*p, d, q*); so, for example, a model described as (0, 1, 2) means that it contains 0 (zero) autoregressive (*p*) parameters and 2 moving average (*q*) parameters which were computed for the series after it was differenced once (documentation.statsoft.com).

The input series needs to be stationary for ARIMA. It should have a constant mean, variance, and autocorrelation through time. Therefore, the series are first differentiated until they are stationary. Sometimes the series also require log transformation of data to stabilize the variance. The *d* parameter indicates the number of times the series needs to be differenced to achieve stationarity (documentation.statsoft.com). The necessary level of differencing is determined by examining the plot of the data and autocorrelogram. Significant changes in level (strong upward or downward changes) usually require first order non seasonal (lag=1) differencing; strong changes of slope usually require second order non seasonal differencing. When the estimated autocorrelation coefficients decline slowly at longer lags, first order differencing is usually needed (documentation.statsoft.com). Some time series may require little or no differencing, and over differenced series produce less stable coefficient estimates It is needed to decide how many autoregressive (*p*) and moving average (*q*) parameters are necessary to yield an effective but has the fewest parameters and greatest number of degrees of freedom among all models that fit the data model of the process. The number of *p* or *q* parameters very rarely need to be greater than 2 in practice (documentation.statsoft.com).

The ARIMA models may also include a constant in addition to the standard p and q parameters. The interpretation of a constant that is statistically significant depends on the model fitted. The expected value of the constant is m, the mean of the series, when there are no AR parameters in the model. The constant represents the intercept, if there are the AR parameters in the series. If the series is differenced, then the constant represents the mean or intercept of the differenced series. If the series investigated is differenced once, and there are no the AR parameters in the model, then the constant represents the mean of the differenced series, and therefore the linear trend slope of the undifferenced series (documentation.statsoft.com).

4. Methodology

The objective of this paper is to investigate effect of autocorrelation on the performance of the Shewhart individual control charts based on original observations or residual charts. This procedure was applied for the data obtained from moisture and ash contents of coal fines which are produced at the Tunçbilek coal washing plant and used as power plant feeding in Tunçbilek, Turkey. The used data were for the -18+0.5 μm size coal fines, called middling product, which is concentrated by dense medium cyclones. The mixture material containing magnetite as dense medium and coals, that sunk in the first dense medium cyclone are feed to second dense medium cyclone flowsheet, having 1800 g/Mg dense medium, for the final separation of -18+0.5 μm coals from schist. The middling product is obtained from and an overflow of cyclone and schist are separated from the underflow of the cyclone. The middling products are then separated from magnetite, dried and sent to the power plant as a feeding material. Daily moisture and ash content data from January 2011 to October 2011 for middling product were supplied from the plant. Totally 242 measurements were obtained and used in the study.

These data were imported into the statistical software package. Trial version of Statgraphics Centurion XVI and Minitab 16 softwares were used to construct different control charts for statistical analyses. First, individual X and moving range (MR) charts assuming normality and independence of observations and X/MR charts applying the Western electric rules to detect the small shifts were generated to monitor the two quality characteristics. Normality and independence of data were checked by applying statistical methods. Since strong autocorrelations were determined for both quality values, the most suitable time series models to remove the autocorrelation between the observations were selected by applying the ARIMA methods. Then, special cause charts (X and MR charts of residuals obtained from ARIMA models) were also generated. Finally, performance of the Shewhart charts of individual observations is compared with the performance of special cause charts for the two coal quality measurements to control the unusual points.

5. Results and discussions

5.1. Control charts for individual coal quality observations

A major drawback of the Shewhart variable charts is its dependency on normality assumption. First, the normal probability plots were generated. When the distribution of the variable is normal, the observed values show a near-linear distribution in the linear line drawn of the chart. The resulted plots including statistical information are given in Fig. 2. As seen in these plots, observed data distribution are almost linear indicating that both variables obey the normal distribution function. The data of the two coal characteristics have been also checked for normality by applying the Anderson-Darling normality test. The Anderson-Darling values were 0.659 and 0.249 for moisture and ash content of coals, respectively. Also p values of both variables were higher than 0.05, suggesting that the data came from a process with normal distribution. The result indicates that the first normality assumption needed for applying the standard Shewhart control charts was satisfied.

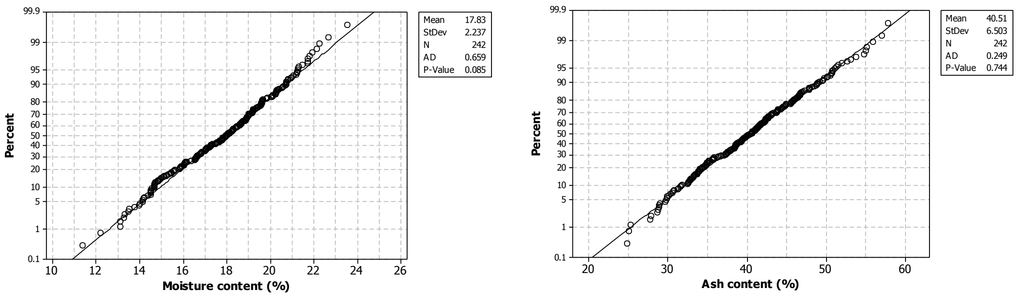

Fig. 2. Normal probability plots of moisture and ash contents

The X and MR control charts of individual data for moisture content and ash contents are given in Fig. 3 and Fig. 4 respectively. The calculated control chart parameters are given in Table 1. The charts are designed to determine whether the data come from a process which is in a state of statistical control. The control charts are constructed under the assumption that the data come from a normal distribution with a mean equal to 17.83 for moisture content, 40.51 for ash content and a standard deviation equal to 1.19 for moisture content, 3.76 for ash content. Of the 242 individual observations shown on the charts, 25 are beyond the 3σ control limits on the X chart while 3 are beyond the limits on the MR chart for moisture content. On the other hand, 20 points are beyond the 3σ control limits on the X chart while 2 points are beyond the limits on the MR chart. Since the probabilities of seeing 25 or more points for moisture content and 20 or more points for ash content beyond the limits just by chance is 0.0. When the data come from the assumed distribution and they are not autocorrelated, we can declare the process to be out of control at the 95% confidence level.

When the Western electric rules were applied to the two coal quality characteristics, much more unusual points were determined in these charts as shown in Figs. 3 and 4. When these plots are examined, we can see that the efficiency of warning limits for detection of process shifting toward out-of-control state is increased when applying the Western Electric rules. Two hundred and 171 unusual runs have been detected for moisture content and ash content, respectively. The plots show observations at which the unusual pattern was detected as well as the particular rule which was violated. For example, rule 3 was violated for point 5 in the moisture content of X chart and for point 18 in the ash content X chart. At these locations, there were groups of 4 out of 5 points beyond 1.0 sigma, all on the same side of the centerline.

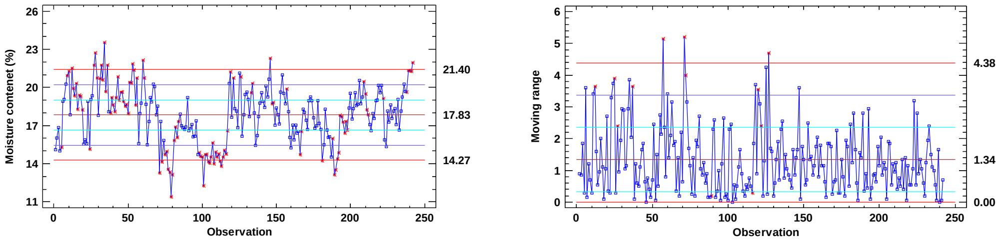

Fig. 3. *X*/MR charts of moisture content and unusual points by Western electric rules

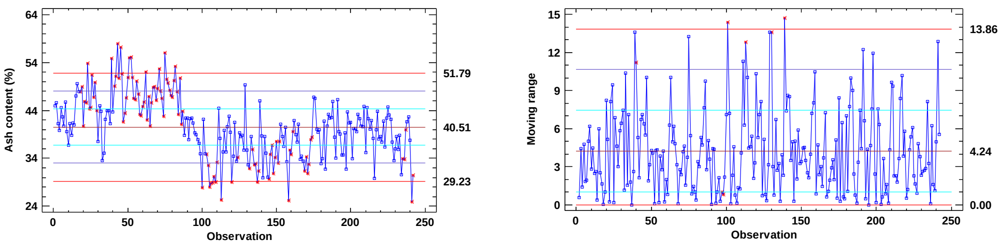

Fig. 3. *X/MR* charts of ash content and unusual points determined by Western electric rules

Table 1. Shewhart chart parameters of moisture and ash contents of coals

X /MR Chart parameters	Moisture content (%)	Ash content (%)
UCL_X	21.39	51.79
$\overline{X}$	17.83	40.51
LCL_X	14.27	29.23
UCL_{MR}	4.38	13.86
$\overline{MR}$	1.34	4.24
UCL_{MR}	0.0	0.0
$1\sigma = \overline{MR}/d_2$	1.19	3.76
2σ	2.37	7.52
3σ	3.56	11.28

Since the normality assumptions are satisfied by the normality tests, the performance of the individual control charts which are discussed so far is valid when the moisture and ash contents of coals are independent. However, if the data are autocorrelated, these findings can result in many wrong decisions about the process. Therefore, whether there is any autocorrelation between the consecutive observations of the two quality characteristics, or not, is tested in the following section.

5.2. Testing autocorrelation of coal characteristics

In addition to normality, independence of data must also be satisfied in order to apply SPC for the processes. Autocorrelation means the correlation between data observed, builds up automatically. If two variables are related in such a way that a change in one variable is reflected by a change in the other variable, the two variables are said to be correlated (Stapenhurst, 2005). Data are autocorrelated if each value is correlated to the previous one. Autocorrelation is a problem because the control charts, and indeed most statistical analyses, assume that the data are not autocorrelated. A quick and simple method of checking for autocorrelation is to draw a scatter diagram of each value, X_i, against the previous value $X_{i\text{-}1}$ (Stapenhurst, 2005). Figure 5 shows the scatter diagram of moisture and ash contents of fine coals used in this study. It is clear that there are considerable positive correlations between the two consecutive data values of both variables. In the autocorrelated series of observation, each individual observation is dependent upon previous observation (Singh and Prajapati, 2011). This also exists for the two coal quality values and the degree of autocorrelation should be determined and also these autocorrelations must be removed by an appropriate ARIMA time series model before constructing statistical control charts.

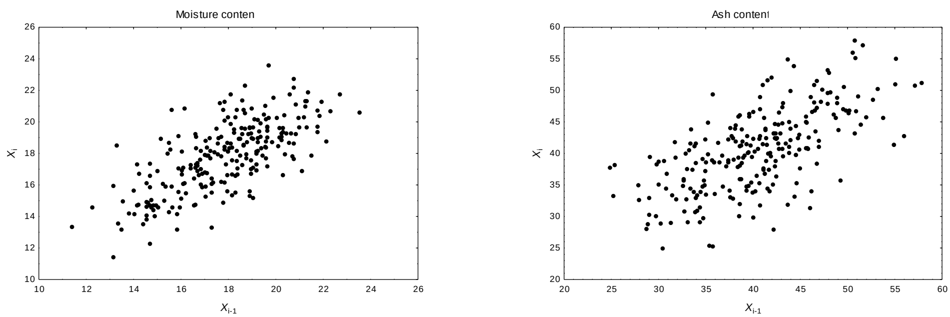

Fig. 5. Scatter diagrams showing autocorrelation for moisture and ash contents

It is possible to calculate the autocorrelation between consecutive values, that is, of lag 1. It is also possible to calculate the autocorrelation of lag 2 (i.e. between values two observations apart) or any number of observations apart (Stapenhurst, 2005). The correlation coefficient can be calculated for lags 1, 2, 3, etc. and it should steadily decrease. For some value *n*, the correlation coefficient will not be significant and this gives us the minimum time that should be allowed between sampling.

To check this, autocorrelation and partial autocorrelation function plots were generated for 25 lags. The sample Autocorrelation Function (ACF) and the sample Partial Autocorrelation Function (PACF) which are representative of the autocorrelation structure across periods of the time series, displayed in Fig. 6 for the moisture content time series (top plots) which show relatively stationary characteristics and for the ash content time series (bottom plots) which show some characteristics of a non-stationary process. There is a high degree of correlation (r = 0.7) for moisture content and (r = 0.6) for ash contents of consecutive data points. The sample ACF of moisture content decays linearly and the corresponding PACF indicates the most important contribution at the first lag. The sample ACF of ash content decays very slowly and the corresponding PACF shows one very significant contribution at lag 1 and then two contributions at lags 2 and 3.

Due to these facts it is likely to see an increased number of false alarms in both Shewhart control charts. Continuing to use a control chart that signals too often can be counter-productive because a real signal may be ignored as just another false alarm.

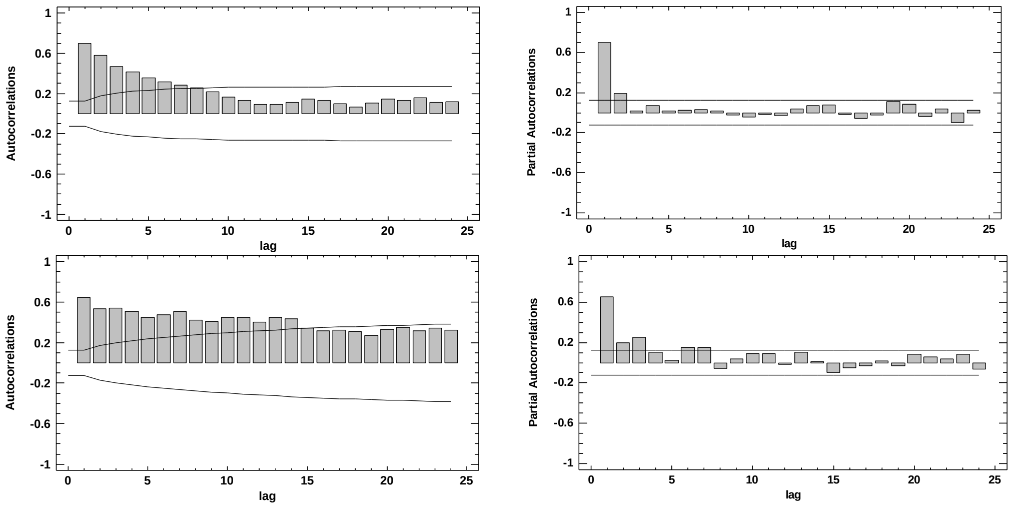

Fig. 6. ACFs and PACFs of moisture (top plots) and ash (bottom plots) contents

5.3. ARIMA fits and residuals

To remove the autocorrelations of two quality characteristics, ARIMA time series models that fit the best to the moisture and ash contents were determined by Statgraphics software. We use the ARIMA procedure in Statgraphics software to estimate the models for moisture and ash contents.

Tables 2 and 3 show the values of model fit statistics used for the ARIMA model selection for the moisture content and ash content respectively. In these tables, A, B, C, D and E denote the ARIMA models that can be suitable for the two quality characteristics. The five ARIMA models in the list are those that fit best, among dozens that were fit. The most suitable model between them was selected by applying the selection criteria. The procedure fits each of the models indicated and selects the models that give the smallest value of the selected criterion. There are six criteria to choose from: (1) the root mean squared error (RMSE), (2) the mean absolute error (MAE), (3) the mean absolute percentage error (MAPE), (4) the mean error (ME), (5) the mean percentage error (MPE), (6) the Akaike Information Criterion (AIC).

Each of the statistics is based on the one-ahead forecast errors, which are the differences between the data value at time t and the forecast of that value made at time t-1. The first three statistics measure the magnitude of the errors. A better model will give a smaller value. The last two statistics measure bias. A better model will give a value close to 0. This table compares the results of fitting different models to the data.

The model with the lowest value of the Akaike Information Criterion (AIC) was chosen as the best describing model in this study. The AIC is calculated from:

$$AIC = 2\ln(RSME) + 2c/n, \tag{13}$$

where $RSME$ is the root mean squared error during the estimation period, c the number of estimated coefficients in the fitted model, and n the sample size used to fit the model. The AIC is a function of the variance of the model residuals, penalized by the number of estimated parameters. In general, the model will be selected that minimizes the mean squared error without using too many coefficients (relative to the amount of data available).

The models with the lowest value of AIC are ARIMA (1,0,1) with a constant for the moisture content and ARIMA (0,1,2) for ash content though several other models are very similar, as presented in Table 2 and Table 3. Therefore, the ARIMA models are preferred for our data in this research.

Table 2. ARIMA model fit statistics for moisture content

Model	RMSE	MAE	MAPE	ME	MPE	AIC
(A)	1.56223	1.24302	7.09856	0.0137352	-0.713845	0.917022
(B)	1.5634	1.24791	7.1247	0.0160021	-0.705291	0.91852
(C)	1.56514	1.24187	7.09152	0.0129439	-0.715767	0.929005
(D)	1.56633	1.24641	7.11658	0.0119596	-0.722733	0.930528
(E)	1.57798	1.23127	6.99972	0.0646808	-0.372863	0.937088

A: ARIMA(1,0,1) with constant; B: ARIMA(2,0,0) with constant; C: ARIMA(1,0,2) with constant; D: ARIMA(2,0,1) with constant; E: ARIMA(2,1,1)

Table 3. ARIMA model fit statistics for ash content

Model	RMSE	MAE	MAPE	ME	MPE	AIC
(A)	4.62662	3.67199	9.47837	-0.168696	-1.69922	3.08018
(B)	4.63186	3.67325	9.5009	-0.178473	-1.74001	3.08245
(C)	4.63362	3.66381	9.425	-0.00738307	-1.29087	3.09147
(D)	4.618	3.61403	9.35612	-0.10959	-1.65678	3.09298
(E)	4.63713	3.66895	9.46142	-0.16297	-1.67918	3.09299

A: ARIMA(0,1,2); B: ARIMA(1,1,1); C: ARIMA(0,1,2) with constant; D: ARIMA(1,0,2) with constant; E: ARIMA(1,1,2)

Parameters of ARIMA (1,0,1) model for moisture and ARIMA (0,1,2) model for ash content are given in Table 4. The output summarizes the statistical significance of the terms of the forcasting model. The terms with *p*-values less than 0.05 are statistically significantly different from zero at the 95.0% confidence level. For the ARIMA (1,0,1) model, the *p*-value for the AR(1) term is less than 0.05, so it is significantly different from 0. The *p*-value for the MA(1) term is less than 0.05, so it is significantly different from 0. The *p*-value for the constant term is less than 0.05, so it is significantly different from 0. The estimated standard deviation of the input which noise is equal to 1.56364. As the evidence of Fig. 6, ACF of the time series values falls down fairly quickly for the moisture content, then the time series values can be considered stationary. Therefore, the moisture values were not need to be differentiated. For the ARIMA (0,1,2) model, the *p*-value for the MA (1) and MA(2) terms are less than 0.05, so they are significantly different from 0. The estimated standard deviation of the input which noise is equal to 4.62806. This model does not contain a constant. As seen in ACF of ash content it goes down very slowly, the time series values of ash content were considered non-stationary (Fig. 6). The ARIMA (0,1,2) model means that the data are differenced once and a second-order moving average term is included, with no intercept.

We obtained the moisture content model in the form:

$X_t = 2.709 + 0.849X_{t-1} + 0.284\varepsilon_{t-1}$ with the AIC of 0.917.

We also obtained the ash content model in the form:

$X_t = X_{t-1} + 0.574\varepsilon_{t-1} + 0.215\varepsilon_{t-2}$ with the AIC of 3.08.

Table 4. Summaries of parameters of ARIMA models

ARIMA (1,0,1) with constant model for moisture content				
Parameter	Estimate	Standard Error	t-value	p-value
AR(1)	0.848526	0.0504203	16.829	0.000000
MA(1)	0.284025	0.0867921	3.27247	0.001224
Mean	17.8889	0.446367	40.0767	0.000000
Constant	2.70971			
ARIMA (0,1,2) model for ash content				
Parameter	Estimate	Standard Error	t-value	p-value
MA(1)	0.574255	0.0627454	9.15215	0.000000
MA(2)	0.215334	0.0633602	3.39857	0.000794

Figure 7 shows the actual and fitted plots of individual observations comparatively. As seen in these figures, the fitted values by ARIMA models were very close to the actual quality characteristics for both variables.

Figure 8 shows the residual autocorrelations and residual partial autocorrelations for moisture and ash contents respectively. As seen in these plots, the autocorrelation removed successfully by applying the ARIMA models and there are no autocorrelations between residual values of two coal quality characteristics. Also, as shown in Fig. 9, the residuals obtained from ARIMA models obey the normal distribution very well. Since the normality and independence assumptions were satisfied by the ARIMA model residuals, the standard Shewhart charts can be applied conveniently to monitor the moisture and ash contents.

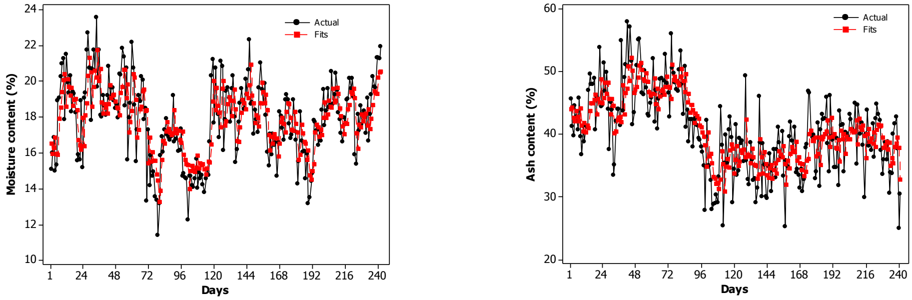

Fig. 7. Actual and fitted moisture contents by ARIMA (1, 0, 1) with constant model and ash contents by ARIMA (0,1,2) model

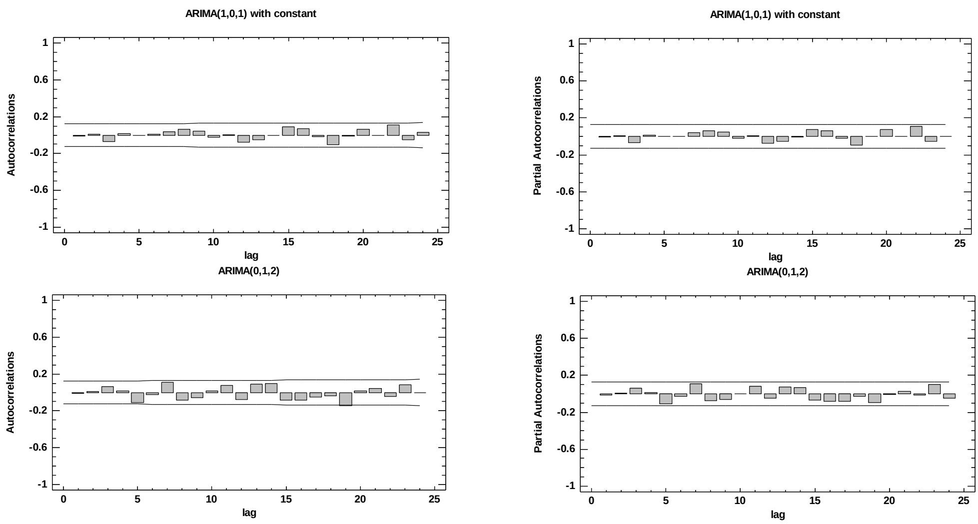

Fig. 8. ACFs and PACFs of residuals for moisture content by ARIMA (1,0,1) with constant model and for ash content by ARIMA (0,1,2)

5.4. X/MR Charts for monitoring forecast residuals

Figure 10 shows the X and MR charts of residuals for individual moisture and ash contents. Even applying the Western electric rules, the number of out-of control points signaficantly reduced the false alarms. There are only 5 and 6 unusual points in X charts of moisture and ash contents, respectively. There is only one point exceeding 3σ limits for moisture and no out-of control point for ash content. On the other hand these numbers are 2 and 3 on the MR charts for the moisture and ash contents, respectively.

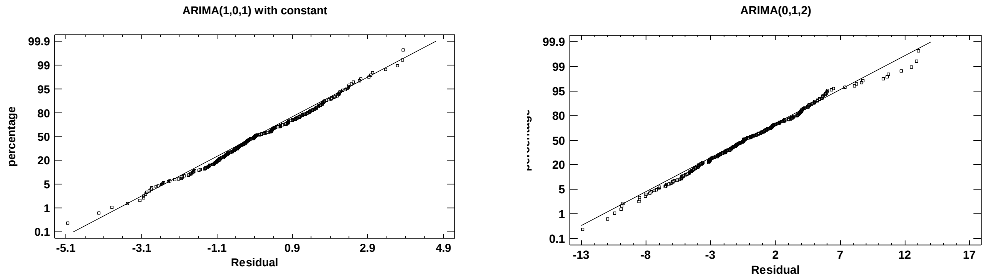

Fig. 9. Normal probability plots of residuals obtained by ARIMA models for moisture and ash contents

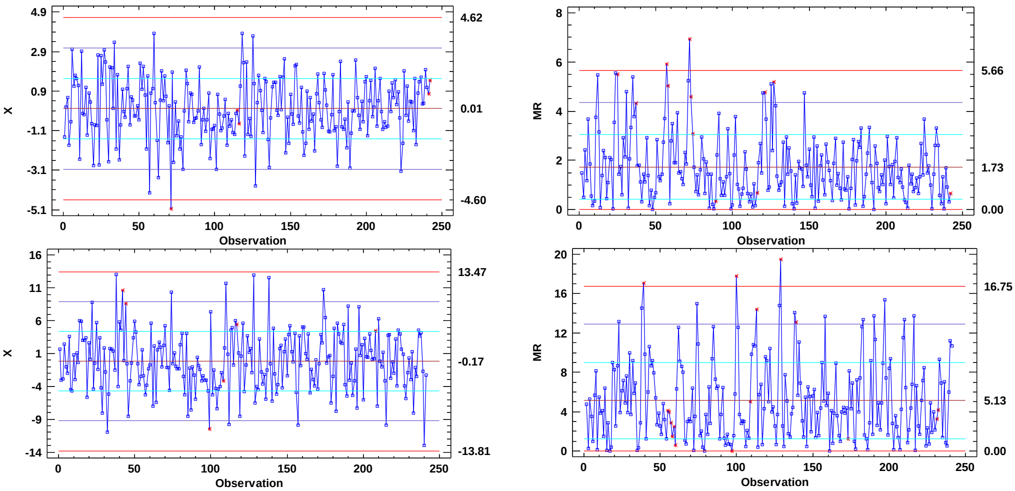

Fig. 10. The *X* and MR charts of residuals (special cause charts) for moisture (top plots) and ash (bottom plots) contents

It is known from literature that these charts would immediately lack the robustness when the observations violated the conditions, and the problem was that most industrial processes are continuous and correlated. When the data were highly correlated, the traditional charts would signal the high rate of false alarms. Alwan and Roberts (1995) analyzed sample applications of control charts and showed that about 85% displayed incorrect control limits were due to the autocorrelation in the data. They pointed out that there would be lots of false prediction or missing alarms when the autocorrelations were ignored. According to results obtained from this study, ignoring autocorrelation in coal data also causes false out-of control points in the SPC charts.

6. Conclusions

Obeying the normality is not enough for applying the SPC charts to monitor two discussed here coal quality characteristics. The results of this study also showed that autocorrelation strongly affected performance of the SPC charts in case of coal washing data. It was determined that the data of daily moisture and ash contents of the coal washing plant were moderately autocorrelated. For this reason, much more out-of control points on the traditional Shewhart charts, which assumes the independence of observations with normaly distributed data, were detected for both variables.

The series of daily values of moisture contents are best described by ARIMA (1,0,1), with a constant model. The ARIMA (0,1,2) model is recommended to be used for generating time series of ash contents. Reduced number of out-of control points in the residual control charts showed us the importance of autocorrelation in coal industry and this autocorrelation should be removed before constructing a control chart to monitor the quality characteristics. It is also shown that the Shewhart control charts of residuals which are independent and normally distributed and obtained from ARIMA models are more appropriate to find the true unusual points otherwise huge amount of wrong out-of control points are recorded. These situations might cause wrong decisions to monitor the process since most of the data points which are in control actually detected as out-of control. Therefore, usage of uncorrelated data may cause misleading results about the process.

It may also possible to make make preventive and corrective actions for the quality characteristics of coal by the ARIMA models since these models can forecast for the future using past data of the processes.

Acknowledgements

The author greatefully acknowledges the Tunçbilek coal washing plant for supplying the data used in this research.

References

ALWAN, L.C., ROBERTS, H.V., 1988, *Time-series modeling for statistical process control*, Journal of Business and Economic Statistics 6, 86–95.

ALWAN, L.C., ROBERTS, H.V., 1995, *The problem of misplaced control limits*, Journal of Royal Statistical Societyi Series C 44(3), 269–278.

ANKARA, H., BİLİR, K., 1995; *Kriblaj tesisinde kalite denetimi*, Madencilikte Bilgisayar Uygulamaları Sempozyumu, 235–240 (in Turkish).

ANKARA, H., YEREL, S., 2008, *Use of range control charts in determination of variability in surface parallelisms of plates*, Journal of Scientific & Industrial Research 67, 1078–1082.

APLEY, D.W., LEE, H.C., 2003, *Design of exponentially weighted moving average control charts for autocorrelated processes with model uncertainty*, Technometrics 45(3), 187–198.

ARSLAN, V., 2006, *Dry coal cleaning methods*, Madencilik 45(3): 9–18 (in Turkish).

AYKUL, H., AKÇAKOCA, H., EDİZ, G.,TAKSUK, M., 2005, *Garp linyitleri işletmesi termik santral kömürleri için istatistiksel süreç kontrol analizi*, Türkiye 19, Uluslararası Madencilik Kongresi, 313–321 (in Turkish).

AYKUL, H., EDİZ, İ. G., AKÇAKOCA, H., ERARSLAN, K., TAKSUK, M., DİXON-HARDY, D. W., 2010, *Statitistical process control for coal fired power plant system*, Journal of the Energy Institute 83, 41–47.

BAYAT, O., ARSLAN, V., 2004, *Statistical analysis in Turkish chromite mining*, Scandinavian Journal of Metallurgy 33, 322–327.

BHATTACHERJEE, A., SAMANTA, B., 2002, *Practical issues in the consruction of control charts in mining applications*, The Journal of The South African Institute of Mining and Metalllurgy, 173–180.
CARSON, P. K., YEH, A. B., 2008, *Exponentially weighted moving average (EWMA) control charts for monitoring an analytical process*, Ind. Eng. Chem. Res. 47, 405–411.
CHANG, Y. M., WU, T. L., 2011, *On average run length of control charts for autocorrelated processes*, Methodol. Comput. Appl. Probab. 13, 419–431.
DEMİRKOL, Ş., 2008, *Comparison of control charts for autocorrelated data*, Master Thesis, Graduate School of Natural and Applied Sciences of Dokuz Eylül University.
ELEVLİ, S., BEHDİOĞLU, S., 2006, İstatistiksel Proses Kontrolü Teknikleri İle Kömür Kalitesindeki Değişkenliğin Belirlenmesi, Madencilik 45(3), 19–26 (in Turkish).
ELEVLİ, S., 2006, *Coal quality control with control charts*, Coal Preparation 26(4), 181–199.
ELEVLİ, S., 2009, *Control charts for autocorrelated colemanite data*, Journal of Scientific & Industrial Research 68, 11–17.
HTTP://DOCUMENTATION.STATSOFT.COM/STATISTICAHelp.aspx?path=TimeSeries/TimeSeries/Overview/Arima/ARIMAARIMAMethodology.
İPEK, H., ANKARA, H., ÖZDAĞ, H., *The application of statistical process control*, Minerals Engineering 12(7): 827-835.
KARAOĞLAN, A.D., BAYHAN, G.M., 2011, *Performance comparison of residual control charts for trend stationary first order autoregressive processes*, Gazi University Journal of Science 24(2), 329–33.
LE, K.H., 1998, *A statistical control chart for process with correlated subgroups*, The Korean Communications in Statistics 5(2), 373–381.
LU, C.W., REYNOLDS, M.R., 2001, *CUSUM charts for monitoring an autocorrelated process*, Journal of Quality Technology 33, 316–334.
MACCHARTY, B.L., WASUSRI, T., 2002, *A review of non-standard applications of statistical process control (SPC) charts*, International Journal of Quality & Reliability Management 19(3), 295–320.
MONTGOMERY, D.C., JENNINGS, C.L., KULAHÇI, M., 2008, *Introduction to time series analysis and forecasting*, Wiley Sries in Probability and Statistics.
MONTGOMERY, D. C., RUNGER, G.C., 2011, *Applied statistics and probability for engineers*, Fifth Edition, John Wiley&Sons, Inc.
SAMANTA, B., BHATTACHERJEE, A., 2001, *An investigation of quality control charts for autocorrelated data*, Mineral Resources Engineering 10, 53–69.
SAMANTA, B., 2002, *Multivariate control charts for grade control using principal-component analysis and time series modelling*, Trans. Inst. Min. Metall. (Sect A: Min Technol) 307, 149–157.
SCHMID, W., SCHÖNE, A., 1997, *Some properties of the EWMA control chart in the presence of autocorrelation*, The Annals of Statistics 25(3): 1277–1283.
STAPENHURST, T., 2005, *Mastering statistical process control*, A handbook for performance improvement using cases, Elsevier.
TEKİR, U., ARSLAN, V., AKAR, G., 2009, *Dry beneficiation of fine coals from soma basin with pneumatic air table*, The Jornal of Ore Dressing 11(22), 33–38.
VAPUR, H., BAYAT, O., AKYOL, F., 2005, *Eti Gümüş A.Ş. Liç Prosesinde İstatistiksel Proses Kontrolü Uygulaması*, Türkiye 19. Uluslararası Madencilik Kongresi, 347-356 (in Turkish).
VAPUR, H., 2009, *Applicability of statistical process control for flotation of metals complex*, Asian Journal of Chemistry 21(1), 655–665.
YEREL, S., ANKARA, H., KONUK, A., ÖZDAĞ, H., 2007, *Preventive maintance policy supported by a quality control chart and Kolmogorov-Smirnov tests: Emet colemanite mineral processing plant, Turkey*, Minerals & Metallurgical Process 24, 152–156.
YEREL, S., ANKARA, H., 2011, *Process control for a coal washing plant using a range control chart and multidimensional, scaling analysis*, Energy Sources, Part A 33, 1028–1034.