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

$$CL = \overline{e}\,(\cong 0), \tag{7}$$

$$UCL = \overline{e} + \frac{3}{1{,}128}\overline{R}_{kl}, \tag{8}$$

$$LCL = \overline{e} - \frac{3}{1{,}128}\overline{R}_{kl}, \tag{9}$$

where

$\overline{e}$ is average value of residuals,

$\overline{R}_{kl}$ is average moving range.

Values CL, UCL and LCL can be calculated as follows

$$CL = \overline{R}_{kl}, \tag{10}$$

$$UCL = 3{,}267 \cdot \overline{R}_{kl}, \tag{11}$$

$$LCL = 0. \tag{12}$$

To increase the sensitivity of control charts ARIMA, it 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 Hotelling chart or CUSUM or EWMA charts for more variables [44].

B. CUSUM control chart

CUSUM (Cumulative Sum) is a sequential analysis technique that is used in the detection of abrupt changes [32].

When applying CUSUM method, a diagram is constructed. On x axis of this diagram a selective order is recorded. On axis y test criterion values Y_k are recorded. The value of test criterion after k-th selective y_k can be defined by formula

$$C_k = \sum_{j=1}^{k}\left(\overline{x}_j - \mu_0\right) = C_{k-1} + \left(\overline{x}_k - \mu_0\right),\ C_0 = 0, \tag{13}$$

where k is the selective order $k = 1, 2, ...$), $\overline{x}_j$ is the selective average from values of regulated value in j-th selection (j = 1, 2, …, k). The study of CUSUM will not answer the question whether the change of diagram progression signals a significant deviation (indicating a definable influence on process) or whether the influence is random. Therefore decision criteria need to be added.

There are two basic types of criteria, through which a decision can be made, whether the process is statistically viable or not. These are:

I. decision mask,

II. decision interval.

These procedures are described in detail in publication [7].

- *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:

On a base of original scale:

$$C_n = \sum_{j=1}^{n}(x_j - \mu)\cdot \tag{14}$$

On a base of normal distribution where the mean μ = 0 and the standard deviation σ = 1:

$$U_j = \frac{(x_j - \mu)}{\sigma},\ S_n = \sum_{j=1}^{n} U_i\ \cdot \tag{15}$$

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 rewritten recurrently [20]: $C_0 = 0$, $C_n = C_{n-1} + (x_n - \mu)$; and with the same principle 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 arbitrary 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 grows linearly with the slope δ. But the population means shift can be more complicated. The CUSUM control chart can reflect all these changes [21].

- *CUSUM for sample means*

We have considered mainly the individual values until now. Now, we will consider subgroups with m observations and calculate the sample means from subgroups. We have to work with the sample mean standard deviation $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{m}}$.

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 $\overline{x}_j$ and the process standard deviation σ with the sample mean standard deviation $\sigma_{\bar{x}}$ in abovementioned formulas [22].

New Process Mean Estimate

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

$$\hat{\mu} = \begin{cases} \mu_0 + K + \dfrac{C_I^+}{N^+} & \text{pro} \quad C_I^+ > H \\[2ex] \mu_0 - K + \dfrac{C_I^-}{N^-} & \text{pro} \quad C_I^- < -H \end{cases}, \tag{16}$$

where N^+ and N^- is a number of selected points from a moment [23], when $C_n^+ = 0$, resp. when $C_n^- = 0$.

- *Comparsion of CUSUM and Shewhart´s Control Charts*

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 Fig. 1) while the Shewhart´s control chart does not detect this deviation [21]. 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 Fig. 1 and Fig. 2) [22].

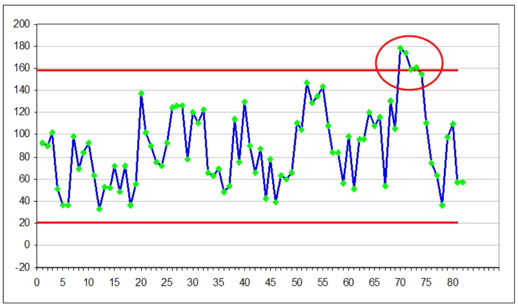

Fig. 1 Shewhart´s Control Chart. Source: QC Expert 2.5Cz

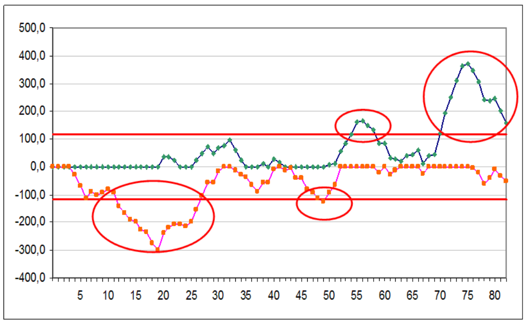

Fig. 2 Control Chart CUSUM. Source: QC Expert 2.5Cz

C. EWMA control charts

Similarly to CUSUM diagram, also EWMA diagram is feasible for situations in which the process is impacted by sudden, small but prevailing changes and the values of observed characteristic are not dependent. Unlike standard diagrams, the regulation bounds depend on the selective moment.

The author of the EWMA diagram is Roberts (1959) in [24]. This diagram works with test criterion Y_k. The value after k-th selection is defined as follows:

$$y_k = (1-\lambda)^k \cdot Y_0 + \lambda \cdot \sum_{j=1}^{k} (1-\lambda)^{k-j} \cdot f(x_j) \qquad (17)$$

for $j = 1, 2, \ldots, k$ and for $0 < \lambda < 1$, $f(x_j)$ is the value of selective characteristic, k is the selection order, Y_0 is the required level of distribution parameter of regulated value.

When dealing with Shewhart regulation diagrams for individual values, we observed that the diagram for individual values is very sensitive to data non-normality in the sense that real, under the ARL (ARL_0) control, would be significantly smaller than the expected value based on the assumption of normal distribution. Borror, Montgomery and Runger (1999) in [25] compared the behavior of ARL Shewhart diagrams for individual values and EWMA diagrams for the case of asymmetric distribution. Features of individuals control charts for Burr distributed and Weibull distributed data see in [31].

In the following text a specific gamma distribution representing the case of asymmetric distribution and t distribution representing symmetrical distribution N(0, 1) will be utilized. ARL_0 values of Shewhart regulation diagrams for individual values and EWMA diagrams for these distributions are denoted in the following tables.

Two aspects in these tables are bewildering. Firstly, even slight non-normality in distribution leads to significant decrease of ARL value in Shewhart diagram for individual values. Subsequently the number of false alarms increases. Secondly EWMA with $\lambda = 0.05$ or $\lambda = 0.10$, and appropriately selected regulation bound can work very well on both symmetric and asymmetric distribution.

With parameters $\lambda = 0.05$ and K = 2.492 the ARL_0 value for EWMA is approximately 8% of the limit emphasized by theory of normal distribution of value $ARL_0 = 370$, with the exception of extreme cases [26].

Tab. 1 ARL values for the EWMA chart and a chart of individual values for different gamma distribution. Source: [26]

	EWMA			Shewhart
λ	**0.05**	**0.1**	**0.2**	**1**
K	**2.492**	**2.703**	**2.86**	**3.00**
Standard	370.4	370.8	370.5	370.4
Gamma (4, 1)	372	341	259	97
Gamma (3, 1)	372	332	238	85
Gamma (2, 1)	372	315	208	71
Gamma (1, 1)	369	274	163	55
Gamma (5, 1)	357	229	131	45

Tab. 2 ARL values for the EWMA chart and a chart of individual values for different *t* distribution. Source: [26]

	EWMA			Shewhart
λ	**0.05**	**0.1**	**0.2**	**1**
K	**2.492**	**2.703**	**2.86**	**3.00**
Standard	370.4	370.8	370.5	370.4
t_{50}	369	365	353	283
t_{40}	369	363	348	266
t_{30}	368	361	341	242
t_8	358	324	259	117
t_6	351	305	229	96
t_4	343	274	188	76

Based on this information the properly designed EWMA is recommended as a control chart for individual values in a wide range of applications, particularly in process monitoring. It is almost a completely nonparametric (independent of distribution) procedure. In addition, EWMA charts are definitely better than Shewhart charts for individual values as well as for the features of the mean shift detection [26], [27].

V. CHANGE POINT DETECTION USING STOCHASTIC DIFFERENTIAL EQUATIONS

Change point estimation consists in the identification of the instant in which a change occurs in the parameter of some model. There are several approaches to the solution of this problem, and here we consider a least squares solution (see, e.g., [4], [5], [8]), but other approaches, such as maximum likelihood change point estimation, are also possible (see, e.g., [28], [34]). We assume we have a diffusion process[1] solution to

$$dX_t = b(X_t)dt + \theta\sigma(X_t)dW_t, \tag{18}$$

where $b(\cdot)$ and $\sigma(\cdot)$ are known functions and $\theta \in \Theta \subset \mathbb{R}$ is the parameter of interest. As in [36], given discrete observations from (18) on $[0, T = n\Delta_n]$, we want to identify retrospectively if and when a change in value of the parameter θ occurred and estimate consistently the parameter before and after the change point. The asymptotics is $\Delta_n \to 0$ as $n \to \infty$ and $n\Delta_n = T$ fixed.[2] For simplicity, we assume that the change occurs at instant k_0, which is one of the integers in 1, …, n. This is a problem of volatility change point estimation that frequently occurs in finance applications. We assume that $\theta = \theta_1$ before the time change and $\theta = \theta_2$ after the time change with $\theta_1 < \theta_2$ (but this does not matter in the final results). In order to obtain a simple least squares estimator, we use Euler approximation. So, from now on, we assume all the hypotheses necessary to have the Euler approximation in place. Namely, we can write the Euler scheme as

$$X_{i+1} = X_i + b(X_i)\Delta_n + \theta\sigma(X_i)(W_{i+1} - W_i)$$

and introduce the standardized residuals

$$Z_i = \frac{(X_{i+1} - X_i) - b(X_i)\Delta_n}{\sqrt{\Delta_n}\,\sigma(X_i)} = \theta\frac{(W_{i+1} - W_i)}{\sqrt{\Delta_n}}.$$

The Z_i's are i.i.d. (independent and identically distributed) Gaussian random variables. The change point estimator is obtained as

$$\hat{k}_0 = \arg\ \min_k \left(\min_{\theta_1,\theta_2} \left\{ \sum_{i=1}^{k} \left(Z_i^2 - \theta_1^2\right)^2 + \sum_{i=k+1}^{n} \left(Z_i^2 - \theta_2^2\right)^2 \right\} \right), \tag{19}$$

with, $k = 2, \ldots, n-1$. We denote by $[x]$ the integer part of the real x, and sometimes we write $k_0 = [n\tau_0]$ and $k = [n\tau]$, $\tau, \tau_0 \in (0,1)$ to indicate the change point in the continuous timescale. Define the partial sums

$$S_n = \sum_{i=1}^{n} Z_i^2, \quad S_k = \sum_{i=1}^{k} Z_i^2, \quad S_{n-k} = \sum_{i=k+1}^{n} Z_i^2,$$

and denote by $\overline{\theta}_1^2$ and $\overline{\theta}_2^2$ the initial least squares estimators of θ_1^2 and θ_2^2 for any given value of k in (19),

$$\overline{\theta}_1^2 = \frac{S_k}{k} = \frac{1}{k}\sum_{i=1}^{k} Z_i^2,$$

and

$$\overline{\theta}_2^2 = \frac{S_{n-k}}{n-k} = \frac{1}{n-k}\sum_{i=k+1}^{n} Z_i^2.$$

These estimators will be refined once a consistent estimator of k_0 is obtained. Denote by U_k^2 the quantity

$$U_k^2 = \sum_{i=1}^{k}\left(Z_i^2 - \overline{\theta}_1^2\right)^2 + \sum_{i=k+1}^{n}\left(Z_i^2 - \overline{\theta}_2^2\right)^2.$$

Then, $\hat{k}_0$ is defined as

$$\hat{k}_0 = \arg\ \min_k U_k^2.$$

To study the asymptotic properties of U_k^2, it is better to rewrite it as

$$U_k^2 = \sum_{i=1}^{n}\left(Z_i^2 - \overline{Z}_n\right)^2 - nV_k^2,$$

where

$$\overline{Z}_n = \frac{1}{n}\sum_{i=1}^{n} Z_i^2$$

and

$$V_k = \left(\frac{k(n-k)}{n^2}\right)^{\frac{1}{2}}\left(\overline{\theta}_2^2 - \overline{\theta}_1^2\right) = \frac{S_n D_k}{\sqrt{k(n-k)}},$$

with

$$D_k = \frac{k}{n} - \frac{S_k}{S_n}.$$

This representations of U_k^2 is obtained by lengthy but straightforward algebra, and it is rather useful because minimization of U_k^2 is equivalent to the maximization of V_k and hence of D_k. So it is easier to consider the following estimator of k_0

$$\hat{k}_0 = \arg\ \max_k |D_k| = \arg\ \max_k \left(k(n-k)\right)^{\frac{1}{2}}|V_k|. \tag{20}$$

As a side remark, it can be noted that, for fixed k (and under suitable hypotheses), D_k is also an approximate likelihood ratio statistic for testing the null hypothesis of no change in volatility (see, e.g., [4]). Once $\hat{k}_0$, has been obtained, the following estimators of the parameters θ_1 and θ_2 can be used:

$$\hat{\theta}_1^2 = \frac{S_{\hat{k}_0}}{\hat{k}_0}, \tag{21}$$

$$\hat{\theta}_2^2 = \frac{S_{n-\hat{k}_0}}{n-\hat{k}_0}. \tag{22}$$

Next results provide consistency of $\hat{k}_0$, $\hat{\theta}_1^2$ and $\hat{\theta}_2^2$ as well as their asymptotic distributions.

[1] For continuous-time observations this problem was studied in [9]. A bayesian approach for discrete-time observations can be found in [35].

[2] For ergodic diffusion processes and $n\Delta_n = T \to \infty$, under additional mild regularity conditions, the results mentioned here are still valid.

Fact 1 ([34]) Under $H_0:\theta_1=\theta_2=1$, we have that

$$\sqrt{\frac{n}{2}}|D_k|\xrightarrow{d}|W^0(\tau)|, \tag{23}$$

where $\{W^0(\tau),0\le\tau\le1\}$ is the Gaussian stochastic process – Brownian bridge, which is useful when we need to model variable, which starts at some point, and that must return to a specific point at a specific time in the future.

The asymptotic result above is useful to test if a change point doesn't exist. In particular, it is possible to obtain the asymptotic critical values for the distribution of the statistic by means of the same arguments used in [34].

Fact 2 ([34]) The estimator $\hat{\tau}_0=\frac{\hat{k}_0}{n}$ satisfies

$$|\hat{\tau}_0-\tau_0|=n^{-1/2}\left(\theta_2^2-\theta_1^2\right)^{-1}O_p\left(\sqrt{\log n}\right). \tag{24}$$

Moreover, for any $\beta\in(0,1/2)$,

$$n^{\beta}\left(\hat{\tau}_0-\tau_0\right)\xrightarrow{p}0.$$

Finally,

$$\hat{\tau}_0-\tau_0=O_p\left(\frac{1}{n\left(\theta_2^2-\theta_1^2\right)^2}\right). \tag{25}$$

It is also interesting to know the asymptotic distribution of $\hat{\tau}_0$ for small discrepancies between θ_1 and θ_2. The case $\vartheta_n=\theta_2^2-\theta_1^2$ equal to a constant is less interesting because when ϑ_n is large the estimate of k_0 is quite precise.

Assumption 1 $\vartheta_n\to0$ in such a way that $\frac{\sqrt{n}\vartheta_n}{\sqrt{\log n}}\to\infty$.

Assumption 1 and Fact 2 imply the consistency of $\hat{\tau}_0$.

Fact 3 ([34]) Under Assumption 1, for $\Delta_n\to0$ as $n\to\infty$, we have that

$$\frac{n\vartheta_n^2\left(\hat{\tau}_0-\tau_0\right)}{2\left(\tilde{\theta}_n^2\right)^2}\xrightarrow{d}\arg\max_v\left\{W(v)-\frac{|v|}{2}\right\}, \tag{26}$$

where $W(u)$ is a two-sided Brownian motion,

$$W(u)=\begin{cases}W_1(-u), & u<0\\ W_2(u), & u\ge0\end{cases}, \tag{27}$$

with W_1 and W_2 two independent Brownian motions and $\tilde{\theta}_n^2$ a consistent estimator for θ_1^2 or θ_2^2.

Finally we have the asymptotic distributions for the estimators $\hat{\theta}_1^2$ and $\hat{\theta}_2^2$ defined in (21) a (22). We denote by θ_0 the common limiting value of θ_1 and θ_2.

Fact 4 ([34]) Under Assumption 1, we have that

$$\sqrt{n}\begin{pmatrix}\hat{\theta}_1^2-\theta_1^2\\ \hat{\theta}_2^2-\theta_2^2\end{pmatrix}\xrightarrow{d}N(0,\Sigma), \tag{28}$$

where

$$\Sigma=\begin{pmatrix}2\tau_0^{-1}\theta_0^4 & 0\\ 0 & 2\left(1-\tau_0\right)^{-1}\theta_0^4\end{pmatrix}. \tag{29}$$

- *Estimation of the change point with unknown drift*

When both $b(x)$ and $\sigma^2(x)$ are unknown, it is necessary to assume that at least $\sigma(x)$ is constant, and hence we consider the stochastic differential equation

$$dX_t=b(X_t)dt+\theta dW_t. \tag{30}$$

Then $b(x)$ can be estimated nonparametrically with

$$\hat{b}_n(x)=\frac{\sum_{i=0}^{n-1}K\left(\frac{x-X_i}{h_n}\right)\left(X_{i+1}-X_i\right)}{\Delta_n\sum_{i=0}^{n-1}K\left(\frac{x-X_i}{h_n}\right)} \tag{31}$$

and Z_i are estimated as

$$\hat{Z}_i=\frac{X_{i+1}-X_i}{\sqrt{\Delta_n}}-\hat{b}_n(X_i)\sqrt{\Delta_n}.$$

Remark 1 In Stanton's approach [34], the two estimators are nothing but Nadaraya-Watson kernel regression estimators of the following conditional expectations

$$b(x)=\lim_{t\to0}\frac{1}{t}\mathrm{E}\left\{X_t-x\mid X_0=x\right\},$$

$$\sigma^2(x)=\lim_{t\to0}\frac{1}{t}\mathrm{E}\left\{\left(X_t-x\right)^2\mid X_0=x\right\}.$$

In this approach, $b(x)$ and $\sigma^2(x)$ are seen as instantaneous conditional means and variances of the process when $X_0=x$. The two quantities can be rewritten, for fixed Δ_n as

$$b(x)=\frac{1}{\Delta_n}\mathrm{E}\left\{X_{i+1}-X_i\mid X_i=x\right\}+\frac{o(\Delta_n)}{\Delta_n},$$

$$\sigma^2(x)=\frac{1}{\Delta_n}\mathrm{E}\left\{\left(X_{i+1}-X_i\right)^2\mid X_i=x\right\}+\frac{o(\Delta_n)}{\Delta_n}.$$

So if we have estimated residues Z_i, in this case, we can use the following contrast to identify the change point:

$$\tilde{k}_0=\arg\min_k\left\{\sum_{i=1}^{k}\left(\hat{Z}_i^2-\frac{\hat{S}_k}{k}\right)^2+\sum_{i=k+1}^{n}\left(\hat{Z}_i^2-\frac{\hat{S}_{n-k}}{n-k}\right)^2\right\}, \tag{32}$$

where

$$\hat{S}_k=\sum_{i=1}^{k}\hat{Z}_i^2 \text{ a } \hat{S}_{n-k}=\sum_{i=k+1}^{n}\hat{Z}_i^2.$$

We obtain the new statistic

$$\hat{V}_k=\left(\frac{k(n-k)}{n^2}\right)^{\frac{1}{2}}\left(\frac{\hat{S}_{n-k}}{n-k}-\frac{\hat{S}_k}{k}\right)=\frac{\hat{S}_n\hat{D}_k}{\sqrt{k(n-k)}},$$

where

$$\hat{D}_k=\frac{k}{n}-\frac{\hat{S}_k}{\hat{S}_n},$$

and the change point is identified as the solution to

$$\hat{k}_0 = \arg \max_k \left|\hat{D}_k\right|.$$

Consistency and distributional results mentioned in [34], [37] and [38].

This paper will be dealt with a change point problem for the volatility of a process solution to a stochastic differential equation, when observations are collected at discrete times. The instant of the change in volatility regime is identified retrospectively by maximum likelihood method on the approximated likelihood. For continuous time observations of diffusion processes Lee, Nishiyama and Yoshida (2006) in [9] considered the change point estimation problem for the drift. In the present work, there will be only assume regularity conditions on the drift process.

VI. Problem Solution

Most data analysis procedures and the resulting conclusions are dependent on fulfilling basic assumptions on which these procedures were based. If not met, all other standard procedures such as calculating an average, confidence intervals, percentiles, most tests, classical Shewhart control charts, design of models for the description of time series, etc., are questionable and impugnable. They usually provide incorrect results and conclusions. A typical breach of conditions for application of control by Shewhart charts or different technologies, but also for the construction of models describing time series, is shown in [41]. The assumptions must be verified by statistical tests, which involve the *basic assumptions for statistical process control:*

- normal data distribution, symmetry,
- constant mean of the process,
- constant variance (standard deviation) of data,
- independence, non-correlation of data,
- absence of outliers and extreme values.

If these assumptions are infringed, conventional regression methods for the time series analysis provide biased and often incorrect results.

Let us now analyze the time series of O/N rate of the Slovak crown for the period 2000 – 2008 (used in [43]). For this purpose the ARIMA model will be used, which is applied if a resulting process is showing such autocorrelation and partial autocorrelation after the transformation of the integrated process using differentiation of d-th order that it is expressed in the form of stationary and invertible ARMA model (p,q). The course of time series is illustrated in the following figure. The graph shows that the time series is non-stationary, but it is not clear whether it contains a seasonal component.

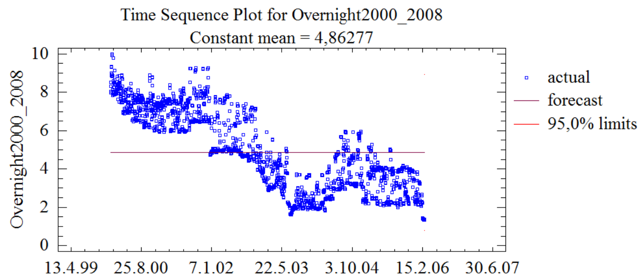

Fig. 3 Time series of O/N rate of the Slovak crown for the period 2000 – 2008. Source: Own Processing

Non-stationarity of the time series is also confirmed by the shape of the ACF and PACF. ACF values fall very slowly and the first value, as well as the PACF, is close to one. The periodogram has a significant peak in the zero (non-seasonal) frequency. Seasonality indicates neither the ACF and PACF, or the periodogram.

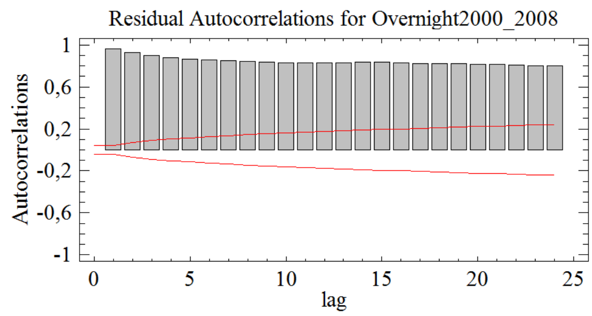

Fig. 4 ACF of the original time series. Source: Own Processing

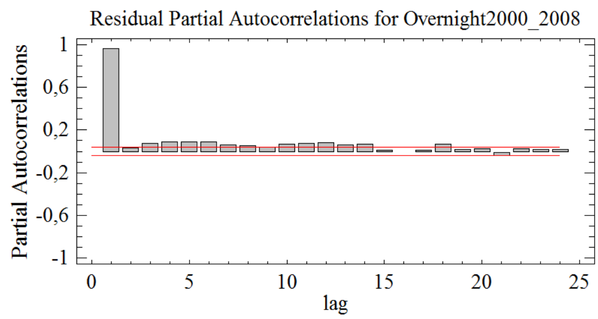

Fig. 5 PACF of the original time series. Source: Own Processing

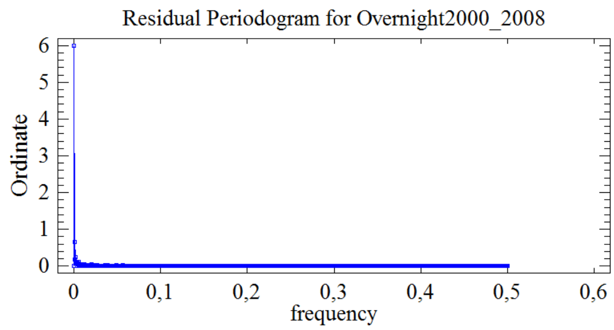

Fig. 6 Periodogram of the original time series. Source: Own Processing

The time series will be again stationarized by the I. non-seasonal difference. Let us skip the elementary steps of the analysis and go directly to the analysis of two ideal models to describe this time series. The first model is after the power linearization ARIMA $(1,1,2)_c$ and the second model is after the logarithmic linearization of the original time series SARIMA$(1,1,2)(1,1,1)_c20$.

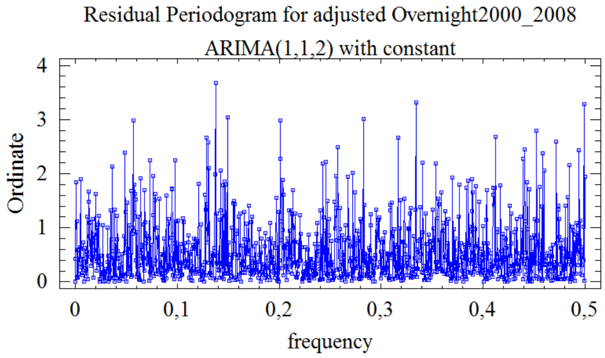

Fig. 7 Residual periodogram for ARIMA model $(1,1,2)_c$. Source: Own Processing

Residual periodogram for ARIMA model $(1,1,2)_c$ shows that residuals are stationary. Now the focus is on the extended SARIMA model (1,1,2) $(1,1,1)_c20$, followed by tables of estimates and interpolation criteria and the very estimates of the model's parameters.

Tab. 3 Estimates of the interpolation criteria of SARIMA model $(1,1,2)(1,1,1)_c20$. Source: Own Processing

	Estimation	*Validation*
Statistic	*Period*	*Period*
RMSE	0.536932	0.145012
MAE	0.327448	0.140548
MAPE	7.23435	10.2761
ME	0.0380698	-0.140548
MPE	-0.398906	-10.2761

Tab. 4 Estimates of parameters of SARIMA model $(1,1,2)(1,1,1)_c20$. Source: Own Processing

Parameter	*Estimate*	*Stand. Error*	*T*	*P-value*
AR(1)	0.752851	0.021549	34.9367	0.000000
MA(1)	0.850656	0.0291219	29.2102	0.000000
MA(2)	0.112643	0.0264917	4.25202	0.000021
SAR(1)	0.046946	0.021776	2.15587	0.031094
SMA(1)	0.974709	0.001605	607.074	0.000000
Mean	-0.00000	0.0000172	-0.420	0.674417
Constant	-0.00000			

Estimated white noise variance = 0.0143047
Estimated white noise standard deviation = 0.1196
Box-Pierce Test
Test based on first 24 autocorrelations
Large sample test statistic = 23.3012
P-value = 0.224336, AIC = -1.25429

Interpolation criteria shows SARIMA $(1,1,2)(1,1,1)_c20$ as much more suitable than ARIMA $(1,1,2)_c$ for the description of this time series.

The figures below show graphs with the time series forecast, autocorrelation and partial autocorrelation functions of unsystematic component for estimated model and residual periodogram.

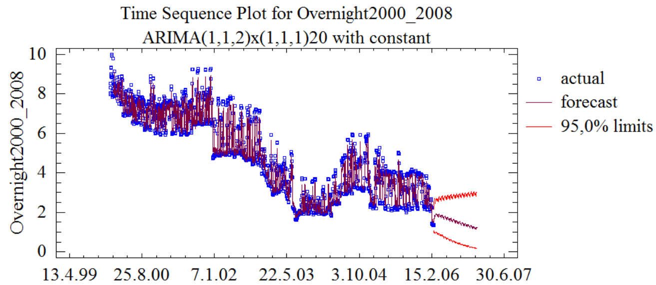

Fig. 8 Graph of the time series with forecast and confidence interval for forecast. Source: Own Processing

According to Akaike information criteria, the Box-Pierce test of autocorrelation of unsystematic component and interpolation criteria, SARIMA model $(1,1,2)(1,1,1)_c20$ appears to be better for the description of this time series than ARIMA model $(1,1,2)_c$.

The following figure illustrates the residual ACF and PACF of the estimated model. P-value of the Box-Pierce test and both of these graphs indicate non-autocorrelation of unsystematic component, thus the estimated model appears to be correct.

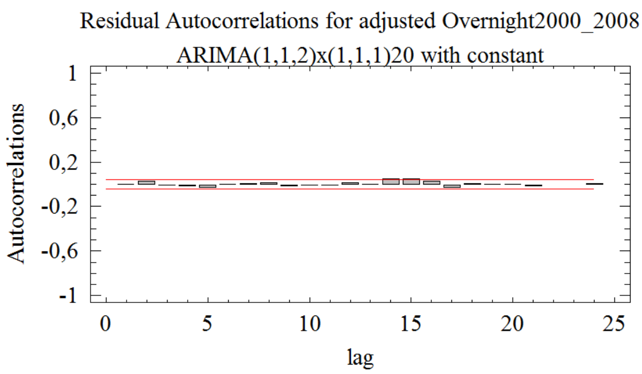

Fig. 9 Residual ACF of SARIMA model $(1,1,2)(1,1,1)_c20$. Source: Own Processing

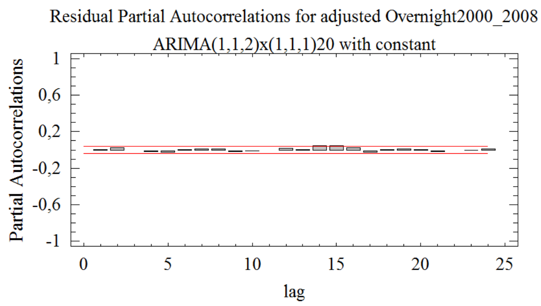

Fig. 10 Residual PACF of SARIMA model (1,1,2)(1,1,1)$_c$20. Source: Own Processing

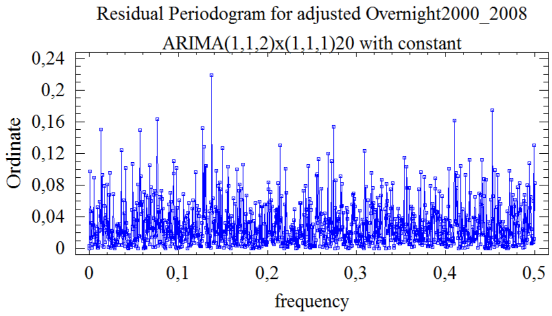

Fig. 11 Residual periodogram of SARIMA model (1,1,2)(1,1,1)$_c$20. Source: Own Processing

Now that we have information on the process due to a statistical model, it can be used to construct control charts for detecting changes in the mean. This detection will be demonstrated on the time series of Overnight values of the Slovak crown during 20 January 2000 – 16 June 2004.

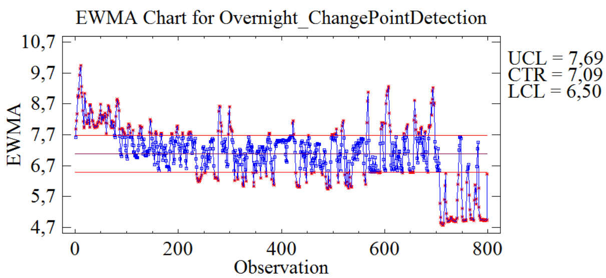

Fig. 12 EWMA control chart for mean shift detection with λ = 0.6. Source: Own Processing

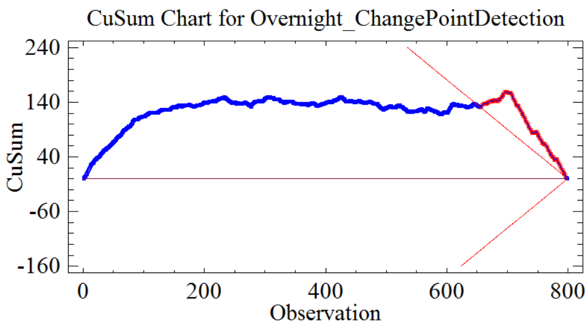

Fig. 13 CUSUM control chart for mean shift detection in 6σ. Source: Own Processing

As evident from the previous images, the EWMA control chart and CUSUM were able to detect changes in the mean almost immediately (EWMA: 13 November 2002; CUSUM: 4 September 2002). These are control charts with memory, therefore information on variability in the time series of SARIMA model (1,1,2)(1,1,1)$_c$20 was used for parameter estimation to construct these diagrams very effectively. The ARIMA control chart detected a shift of mean and therefore high heteroscedasticity in 5 August 2002, which corresponds to fluctuations of Slovak crown.

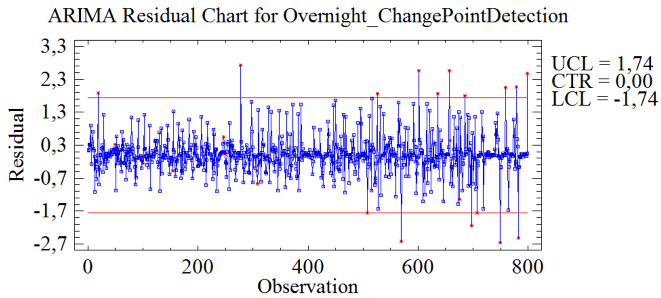

Fig. 14 ARIMA control chart for volatility change point detection. Source: Own Processing

Then we look at the volatility change point detection using the SDEs least squares approach.

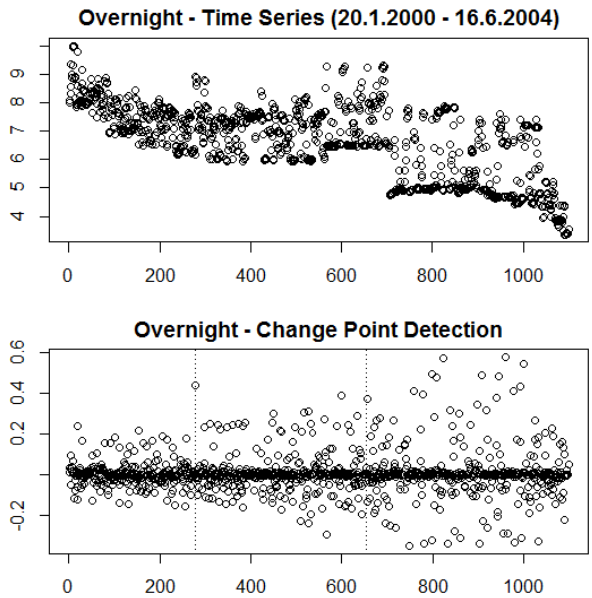

Fig. 15 Overnight values change point analysis of Overnight values of the Slovak crown in the period 20 January 2000 – 16 June 2004. Source: Own Processing

All the analysis using SDEs have been done using the R statistical environment (R Development Core Team, 2009) and the package `sde` (see in [37]) and Yuima (see in [42]).

Following is the output of the R programming language environment using the `sde` package.

```
$k0
[1] 655
$tau0
[1] 655 (2 September 2002)
$theta1
[1] 0.07893081
$theta2
[1] 0.124977
```

Looking at the previous figure, another change point may be present. So we reanalyze the first part of the series to spot the second change point.

```
$k0
[1] 276
$tau0
[1] 276 (23 February 2001)
$theta1
[1] 0.05829864
$theta2
[1] 0.08956524
```

Both change points (23 February 2001 and 2 September 2002) correspond to fluctuations of Slovak crown.

VII. 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.

As for the two sub-series identified by the change point estimate and estimators of these parameters: both are consistent and asymptotically normal at the usual rate $\sqrt{n}$, with n the number of observations. The least squares estimator $\hat{\tau}_0$ seems to have a good performance in terms of bias and variability for models with constant or bounded drift, while it behaves badly in the presence of unbounded drift as time T grows.

VIII. Conclusion

In financial markets it is crucial to have an accurate description of the volatility of the market and/or the different financial products. All pricing formulas make use of the historical values of the volatility as a fundamental ingredient. It is well known, however, that volatility is not constant over time, even short time, and thus the monitoring of the volatility is one of the primary tasks in empirical finance. The statistical way to monitor structural changes is called change point analysis or change point estimation.

The paper dealt with the control charts and SDEs 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 financial flows. 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 upgradeable to handle SPC functions. Under such an integrated scheme, the usefulness of the proposed procedure will be optimized. Based on information from chapter 4, we would recommend a properly designed time series control charts as control charts for individual measurements in a wide range of applications. They are almost perfectly nonparametric (distribution-free) procedures.

Stochastic differential equations are among the most used stochastic models to describe continuous time financial time series. Although data are collected in discrete time, the underlying structure of the continuous model allows for very detailed analysis of these data. In the analysis of the Slovak crown, the approach revealed (unlike time series control charts) another change point in the structure of variability, thus appearing to be preferable for the change point detection. It also shows that the use of SDEs provides robustness in the estimation results.

References

[1] F. Black and M. S. Scholes, "The Pricing of Options and Corporate Liabilities (Book style)," *Journal of Political Economy*, vol. 81, 1973, pp. 637–54.

[2] R. C. Merton, "Theory of Rational Option Pricing (Book style)," *Bell Journal of Economics and Management Science*, vol. 41, 1974, pp. 141–83.

[3] D. V. Hinkley, "Inference about the change-point from cumulative sum tests," *Biometrika*, vol. 58, 1971, pp. 509–23.

[4] C. Inclan and G. C. Tiao, "Use of Cumulative Sums of Squares for Retrospective Detection of Change of Variance," *Journal of the American Statistical Association*, vol. 89, 1994, pp. 913–23.
[5] J. Bai, "Least Squares Estimation of a Shift in Linear Processes," *Journal of Times Series Analysis*, vol. 15, 1994, pp. 453–72.
[6] S. Kim, S. Cho and S. Lee, "On the Cusum Test for Parameter Changes in GARCH(1,1) Models," *Communications in Statistics - Theory and Methods*, vol. 29, 2000, pp. 445–62.
[7] S. Lee, J. Ha, O. Na and S. Na, "The Cusum Test for Parameter Change in Time Series Models," *Scandinavian Journal of Statistics*, vol. 30, 2003, pp. 781–96.
[8] G. Chen, Y. K. Choi and Y. Zhou, "Nonparametric Estimation of Structural Change Points in Volatility Models for Time Series," *Journal of Econometrics*, vol. 126, 2005, pp. 79–144.
[9] S. Lee, Y. Nishiyama, N. Yoshida, "Test for Parameter Change in Diffusion Processes by Cusum Statistics Based on One-Step Estimators," *Annals of the Institute of Statistical Mathematics*, vol. 58, 2006, pp. 211–22.
[10] M. Basseville and I. Nikiforov, *Detection of abrupt changes: theory and application*, Englewood Cliffs, Ed. New York: Prentice Hall, 1993, ch. 2.
[11] D. Nazario, B. Ramirez and S. Tep, "Transient detection with an application to a chemical process," *Computers ind. Eng*, 1997, pp. 896-908.
[12] R. B. Blazek, K. HongJoong, R. Boris and T. Alexander, "A novel approach to detection of "denial–of–service" attacks via adaptive sequential and batch-sequential change-point detection methods," *in Proceedings of the 2001 IEEE*, 2001, pp. 220-226.
[13] R. Lund, L. W. Xiaolan, Q. Lu, J. Reeves, C. Gallagher and Y. Feng, "Change-point detection in periodic and autocorrelated time series," *Journal of Climate* (20), 2007, pp. 5178-5190.
[14] V. Moskvina and A. Zhigljavsky, "An algorithm based on singular spectrum analysis for changepoint detection," *Communication in Statistics, Simulation and Computation*, vol. 32, no. 2, 2003, pp. 319-352.
[15] M. Mboup, C. Join and Ml. Fliess, "An online change-point detection method," *in Proceedings of the Mediterranean Conference on Control and Automation Congress Centre,* Ajaccio, France: IEEE, 2008, pp. 1290-1295.
[16] L. Auret and C. Aldrich, "Unsupervised process fault detection with random forests. Industrial and Engineering Chemistry Research," vol. 49(19), 2010, pp. 9184-9194.
[17] X. Wei, H. Hanping, Y. Yue and Q. Wang, "Anomaly detection of network traffic based on the largest Lyapunov exponent," *IEEE*, 2010, pp. 581-585.
[18] C. Aldrich and G. T. Jemwa, "Detecting change in complex process systems with phase space methods," *The 2007 Spring National Meeting, 2007 – AIChE*, 2007, pp. 1-10.
[19] L. A. Vincent, "A technique for the identification of inhomogeneities in Canadian temperature series," *Journal of Climate Change*, 1998, pp. 1094-1104.
[20] M. J. Chandra, *Statistical Quality Control*. United States of America: CRC Press, LLC, 2001, ch. 8.
[21] T. J. Harris and W. H. Ross, "Statistical process control procedures for correlated observations," *Canadian Journal of Chemical Engineering*, vol. 69, 1991, pp. 48-57.
[22] C. W. Lu and M. R. Reynolds, "EWMA control charts for monitoring the mean of autocorrelated processes," *Journal of Quality Technology*, vol. 31, 1999, pp. 166-188.
[23] D. S. Chambers and D. J. Wheeler, *Understanding Statistical Process Control*, 2nd ed. USA: SPC Press, Inc, 1992, ch. 3.
[24] S. W. Roberts, "Control chart tests Based on Geometric Moving Averages," *Technometrics*, vol. 1, 1959, pp. 239-250.
[25] C. M. Borror, D. C. Montgomery and G. C. Runger, "Robustness of the EWMA. Control Chart to Non-normality," *Journal of Quality Technology*, vol. 31(3), 1999, pp. 309–316.
[26] D. Montgomery, *Introduction to Statistical Quality Control*, 6th edition. New York: John Wiley & Sons, Inc, 2009, ch. 9.
[27] C. W. Lu and M. R. Reynolds, "Control charts for monitoring the mean and variance of autocorrelated processes," *Journal of Quality Technology*, vol. 31, no. 3, 1999, pp. 259-274.
[28] J. Bai, "Estimation of a change point in multiple regression models," *Rev. Econ. Stat.*, vol. 79, 1997, pp. 551–563.
[29] A. S. Jamali and L. JinLin, "False Alarm Rates for the Shewhart Control Chart with Interpretation Rules," *WSEAS Transactions on Information Science and Applications*, vol. 3, 2006, pp. 7.
[30] Ch. Ting and Ch. Su, "Using Concepts of Control Charts to Establish the Index of Evaluation for Test Quality," *WSEAS Transactions on Communications,* vol. 8, 2009, pp. 6.
[31] B. Kan and B. Yazici, "The Individuals Control Charts for Burr Distributed and Weibull Distributed Data," *WSEAS Transactions on Mathematics*, vol. 5, 2006, pp. 5.
[32] Z. M. Nopiah, M. N. Baharin, S. Abdullah, M. I. Khairir and C. K. E. Nizwan, "Abrupt Changes Detection in Fatigue Data Using the Cumulative Sum Method," *WSEAS Transactions on Mathematics*, vol. 7, 2008, pp. 12.
[33] N. K. Huat and H. Midi, "Robust Individuals Control Chart for Change Point Model," *WSEAS Transactions on Mathematics*, vol. 9, 2010, pp. 7.
[34] M. Csörgö and L. Horvath, *Limit Theorems in Change-point Analysis*. New York: Wiley, 1997, ch. 1.
[35] J. Liechty and G. Roberts, "Markov chain Monte Carlo methods for switching diffusion models," *Biometrika*, vol. 88(2), 2001, pp. 299–315.
[36] A. De Gregorio and S. M. Iacus. (2007). Least squares volatility change point estimation for partially observed diffusion processes, Working Paper. Available: http://arxiv.org/abs/0709.2967v1.
[37] S. M. Iacus, *Simulation and Inference for Stochastic Differential Equations: with R Examples*, Springer Series in Statistics, New York: Springer, 2008, ch. 4.
[38] S. M. Iacus and N. Yoshida (2009). Estimation for the Change Point of the Volatility in a Stochastic Differential Equation. Available: http://arxiv.org/abs/0906.3108.
[39] Y. Kutoyants, *Identification of Dynamical Systems with Small Noise*, Dordrecht: Kluwer, 1994, ch. 5.
[40] Y. Kutoyants, *Statistical Inference for Ergodic Diffusion Processes*, London: Springer-Verlag, 2004, ch. 4.
[41] M. Meloun and J. Militky, *Kompendium statistickeho zpracovani dat.* Praha: Academia, Nakladatelstvi Akademie ved Ceske republiky, 2006, ch. 10.
[42] YUIMA Project Team. (2010). Yuima: The YUIMA Project Package. Available: http://R-Forge.R-project.org/projects/yuima.
[43] M. Kral and M. Kovarik, *Carry Trade*, 1st ed., Zilina: Georg, 2011, ch. 6.
[44] D. Noskievicova, "Vybrane metody statisticke regulace procesu pro autokorelovana data," *Automa*, vol.10, no.1, 2008, pp. 40-43.
[45] R. Husek, *Ekonometricka analyza*. Praha: Vysoka skola ekonomicka v Praze, Nakladatelstvi Oeconomica, 2007, ch. 6.

M. Kovarik graduated at the Faculty of Management and Economics, Tomas Bata University in Zlin where he is lecturing at the Department of Statistics and Quantitative Methods since 2009. He also graduated at Faculty of Applied Informatics, Tomas Bata University in Zlin in the field of Information Technology. Author and co-author of 7 books and 5 lecture notes, his research is focused on mathematical and statistical methods in quality management and computationally-intensive statistical data analyses with results published in numerous peer-reviewed journals and presented at conferences in the Czech Republic as well as internationally. Martin Kovářík is also a consultant of statistical data analysis, application of statistical methods in quality management and questionnaire-based surveys data processing. A participant in several successful academic projects, specifically Innovation of Follow-up Master's degree programmes at the Faculty of Management and Economics, reg. no. CZ.1.07/2.2.00/07.036 and IGA's (Internal Grant Agency) Development of mathematical and statistical methods utilization in quality management, reg. no. IGA/73/FaME/10/D, he also cooperates with organizations such as Barum Continental, s. r. o. and Tomas Bata Regional Hospital, a. s.