

# **Návrh monitorovacího aparátu pro autokorelované procesy**

The proposal of monitoring system for autocorrelated processes

Bc. Martin Kovářík

---

Diplomová práce  
2013



Univerzita Tomáše Bati ve Zlíně  
Fakulta aplikované informatiky

---

Univerzita Tomáše Bati ve Zlíně  
Fakulta aplikované informatiky  
akademický rok: 2012/2013

## ZADÁNÍ DIPLOMOVÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Ing. Martin Kovářik**  
Osobní číslo: **A11532**  
Studijní program: **N3902 Inženýrská informatika**  
Studijní obor: **Automatické řízení a informatika**  
Forma studia: **kombinovaná**

Téma práce: **Návrh monitorovacího aparátu pro autokorelované procesy**

Zásady pro vypracování:

1. Vypracujte literární rešerši současného stavu dané problematiky autokorelace v SPC.
2. Na základě literární rešerše porovnejte výkonnost klasických Shewhartových regulačních diagramů a regulačních diagramů pro časové řady v případě autokorelace v datech.
3. Navrhněte způsob řešení typických porušení předpokladů pro aplikaci regulace Shewhartovými diagramy a konstrukci modelů pro popis časových řad.
4. Navrhněte monitorovací aparát pro autokorelované procesy, který odhalí změny ve střední hodnotě, rozptylu a v autokorelační struktuře procesních dat.
5. Simulačním způsobem ověřte stabilitu a výkonnost tohoto navrženého aparátu.

Rozsah diplomové práce:

Rozsah příloh:

Forma zpracování diplomové práce: **tištěná/elektronická**

Seznam odborné literatury:

1. FUCHS, C., KENETT, Ron S., 1998. **Multivariate Quality Control**. 1. vyd. New York: Marcel Dekker, Inc. 212 s. ISBN 0-8247-9939-9.
2. HUŠEK, R., 2007. **Ekonometrická analýza**. 1. vyd. VŠE Praha: Oeconomica. 367 s. ISBN 978-80-245-1300-3.
3. CHAMBERS David S., WHEELER, Donald J., 1992. **Understanding Statistical Process Control**. 2nd edition. USA: SPC Press, Inc. 300 s. ISBN 0-945320-13-2.
4. CHANDRA, M., 2001. **Statistical Quality Control**. 1. vyd. USA: CRC Press, LLC. 280 s. ISBN 0-8493-2347-9.
5. MELOUN, M., MILITKÝ, J., 2006. **Kompendum statistického zpracování dat**. 2. vyd. Praha: Academia, nakladatelství Akademie věd České republiky. 982 s. ISBN 80-200-1396-2.
6. MONTGOMERY, Douglas C., 2009. **Introduction to Statistical Quality Control**. 6. vyd. USA: John Wiley & Sons, Inc. 734 s. ISBN 978-0-470-16992-6.

Vedoucí diplomové práce: **prof. Ing. Vladimír Bobál, CSc.**  
Ústav řízení procesů

Datum zadání diplomové práce: **24. února 2013**

Termín odevzdání diplomové práce: **11. června 2013**

Ve Zlíně dne 24. února 2013

  
prof. Ing. Vladimír Vašek, CSc.  
*děkan*



  
prof. Ing. Vladimír Vašek, CSc.  
*ředitel ústavu*

## ABSTRAKT

Klasické metody statistického řízení procesů (SPC), jako jsou Shewhartovy regulační diagramy a CUSUM regulační diagramy, předpokládají, že získaná procesní data jsou nezávislá. Nicméně, tento předpoklad byl zpochybněn, jakmile bylo zjištěno, že v mnoha praktických situacích, jsou data sériově korelována. Výkonnost klasických regulačních diagramů se v případě autokorelace v datech výrazně zhorší. Cílem této diplomové práce je provést literární rešerši současného stavu dané problematiky autokorelace v SPC, porovnat výkonnost klasických Shewhartových regulačních diagramů a regulačních diagramů pro časové řady v případě autokorelace v datech a navrhnout monitorovací aparát pro autokorelované procesy, který odhalí změny ve střední hodnotě, rozptylu a v autokorelační struktuře.

Klíčová slova: Statistické řízení procesů (SPC), Shewhartův regulační diagram, diagram EWMA, diagram CUSUM, diagram SACC, MH diagram, diagram  $\lambda_{LS,max}$

## ABSTRACT

Traditional Statistical Process Control (SPC) methods such as Shewhart's and CUSUM regulation schemes presuppose processing data's independence. However, the requirement has been found not to hold in practical situations due to serial correlation. Performance of traditional diagrams substantially decreases when dealing with autocorrelated data. The aim of the master thesis is to evaluate existing literature sources dealing with autorrelation in SPC, compare performance of traditional Shewhart's regulation diagrams and diagrams for time-series analyses for autocorrelated data, and devise monitoring framework for autocorrelated processes capable of detecting changes in mean, variance, and autocorrelation structure.

Keywords: Statistical process control (SPC), Shewhart control chart, EWMA chart, CUSUM chart, SACC chart, MH chart,  $\lambda_{LS,max}$  chart

Touto cestou bych především velmi rád poděkoval svojí rodině, která mě po celou dobu studia podporovala a chápala moji občasnou myšlenkovou nepřítomnost v období psaní této práce. Dále bych chtěl poděkovat vedoucímu mé diplomové práce panu prof. Ing. Vladimíru Bobálovi, CSc. za odborné vedení, cenné informace a připomínky, které mi při vypracování diplomové práce poskytl.

V neposlední řadě bych chtěl také poděkovat vedení společnosti Continental Barum, s.r.o. za poskytnutí informací a podkladů pro zpracování této diplomové práce.

**Prohlašuji, že**

- beru na vědomí, že odevzdáním diplomové/bakalářské práce souhlasím se zveřejněním své práce podle zákona č. 111/1998 Sb. o vysokých školách a o změně a doplnění dalších zákonů (zákon o vysokých školách), ve znění pozdějších právních předpisů, bez ohledu na výsledek obhajoby;
- beru na vědomí, že diplomová/bakalářská práce bude uložena v elektronické podobě v univerzitním informačním systému dostupná k prezenčnímu nahlédnutí, že jeden výtisk diplomové/bakalářské práce bude uložen v příruční knihovně Fakulty aplikované informatiky Univerzity Tomáše Bati ve Zlíně a jeden výtisk bude uložen u vedoucího práce;
- byl/a jsem seznámen/a s tím, že na moji diplomovou/bakalářskou práci se plně vztahuje zákon č. 121/2000 Sb. o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon) ve znění pozdějších právních předpisů, zejm. § 35 odst. 3;
- beru na vědomí, že podle § 60 odst. 1 autorského zákona má UTB ve Zlíně právo na uzavření licenční smlouvy o užití školního díla v rozsahu § 12 odst. 4 autorského zákona;
- beru na vědomí, že podle § 60 odst. 2 a 3 autorského zákona mohu užít své dílo – diplomovou/bakalářskou práci nebo poskytnout licenci k jejímu využití jen s předchozím písemným souhlasem Univerzity Tomáše Bati ve Zlíně, která je oprávněna v takovém případě ode mne požadovat přiměřený příspěvek na úhradu nákladů, které byly Univerzitou Tomáše Bati ve Zlíně na vytvoření díla vynaloženy (až do jejich skutečné výše);
- beru na vědomí, že pokud bylo k vypracování diplomové/bakalářské práce využito softwaru poskytnutého Univerzitou Tomáše Bati ve Zlíně nebo jinými subjekty pouze ke studijním a výzkumným účelům (tedy pouze k nekomerčnímu využití), nelze výsledky diplomové/bakalářské práce využít ke komerčním účelům;
- beru na vědomí, že pokud je výstupem diplomové/bakalářské práce jakýkoliv softwarový produkt, považují se za součást práce rovněž i zdrojové kódy, popř. soubory, ze kterých se projekt skládá. Neodevzdání této součásti může být důvodem k neobhájení práce.

**Prohlašuji,**

- že jsem na diplomové práci pracoval samostatně a použitou literaturu jsem citoval. V případě publikace výsledků budu uveden jako spoluautor.
- že odevzdaná verze diplomové práce a verze elektronická nahraná do IS/STAG jsou totožné.

Ve Zlíně

.....  
podpis diplomanta

**OBSAH**

<b>SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK.....</b>	<b>9</b>
<b>ÚVOD.....</b>	<b>11</b>
<b>TEORETICKÁ ČÁST.....</b>	<b>12</b>
<b>1 INTERPRETACE REGULAČNÍHO DIAGRAMU.....</b>	<b>13</b>
1.1 OBECNÝ POSTUP SESTROJENÍ A ANALÝZY REGULAČNÍHO DIAGRAMU.....	13
1.2 KLASICKÉ SHEWHARTOVY REGULAČNÍ DIAGRAMY.....	14
1.3 ÚČINNOST REGULAČNÍHO DIAGRAMU .....	17
1.4 ČLENĚNÍ REGULAČNÍCH DIAGRAMŮ .....	23
1.5 SHEWHARTOVY REGULAČNÍ DIAGRAMY PRO REGULACI MĚŘENÍM .....	24
1.6 VLIV AUTOKORELACE NA KLASICKÉ REGULAČNÍ DIAGRAMY .....	25
<b>2 REGULAČNÍ DIAGRAMY PRO ČASOVÉ ŘADY.....</b>	<b>28</b>
2.1 METODA KUMULOVANÝCH SOUČTŮ (CUSUM) .....	28
2.1.1 Základní charakteristiky metody CUSUM.....	29
2.1.2 Definování diagramu CUSUM pro výběrové průměry ( $\bar{x}$ ) .....	31
2.1.3 Interpretace průběhu diagramu CUSUM .....	31
2.2 DIAGRAMY EWMA .....	32
2.2.1 Klasický diagram EWMA pro výběrové průměry ( $\bar{x}$ ).....	33
2.3 ARIMA A SARIMA PROCESY .....	35
2.3.1 Oblast použití .....	35
2.3.2 Procesy ARIMA.....	35
2.3.3 Sezónní procesy – Procesy SARIMA .....	36
2.3.4 Identifikace a ověřování modelu, konstrukce předpovědí .....	37
2.3.4.1 Identifikace modelu .....	37
2.3.4.2 Diagnostická kontrola modelu .....	39
2.3.4.3 Výpočet předpovědí.....	40
2.4 REGULAČNÍ DIAGRAM ARIMA .....	41
2.4.1 Hledání modelu ARIMA.....	45
<b>PRAKTICKÁ ČÁST .....</b>	<b>46</b>
<b>3 SYSTÉM AUTOMATICKÉHO SBĚRU DAT PRO MONITOROVÁNÍ KVALITY V PODNIKU CONTINENTAL BARUM, S.R.O. – SPC SOFTWARE .....</b>	<b>47</b>
3.1 QTREE-SPC C/S.....	50
3.2 OVLADAČE A EXTERNÍ ZAŘÍZENÍ .....	52
<b>4 ŘEŠENÍ PROBLEMATIKY MONITOROVÁNÍ AUTOKORELOVANÝCH PROCESŮ V SOUČASNÉ PRAXI.....</b>	<b>53</b>
4.1 PŘÍPADOVÁ STUDIE SESTROJENÍ DYNAMICKÉHO DIAGRAMU EWMA.....	53
4.2 PŘÍPADOVÁ STUDIE APLIKACE POUŽITÍ REGULAČNÍHO DIAGRAMU EWMA .....	59
4.3 PRAKTICKÁ PŘÍPADOVÁ STUDIE – POUŽITÍ REGULAČNÍCH DIAGRAMŮ ARIMA, EWMA A CUSUM NA REÁLNÉM PROCESU .....	62

<b>5</b>	<b>NÁVRH REGULAČNÍHO APARÁTU PRO AUTOMATIZOVANÝ SBĚR DAT VE VÝROBNÍCH PROCESECH.....</b>	<b>69</b>
5.1	SIMULTÁNNÍ MONITOROVÁNÍ STŘEDNÍ HODNOTY, ROZPTYLU A AUTOKORELAČNÍ STRUKTURY SÉRIOVĚ KORELOVANÝCH PROCESŮ.....	69
5.1.1	Návrh nového přístupu pro detekci posunu v MVAS.....	71
5.1.1.1	Vlastnosti a provedení ARL .....	73
5.1.1.2	Citlivost na detekci posunu střední hodnoty.....	74
5.1.1.3	Citlivost na posun rozptylu .....	75
5.1.1.4	Citlivost na detekci změn v autokorelační struktuře.....	76
5.1.2	Výpočetní simulační studie .....	77
5.2	NÁVRH APARÁTU PRO MONITOROVÁNÍ AUTOKORELOVANÝCH PROCESŮ.....	79
5.3	ZJIŠŤOVÁNÍ ODLEHLÝCH HODNOT A ÚROVNĚ POSUNU .....	82
5.4	VLASTNOSTI $\lambda_{LS,T}$ .....	86
5.4.1	Efekt úrovně posunu .....	86
5.4.2	Vliv odlehlých hodnot.....	91
5.5	NÁVRH MONITOROVACÍHO APARÁTU .....	92
5.5.1	Implementace regulačního diagramu .....	93
5.5.2	Výkonnost, vlastnosti a chování regulačního diagramu .....	94
5.5.3	Srovnání ARL .....	95
<b>6</b>	<b>ANALÝZA PRO VÝBĚR VHODNÉHO STATISTICKÉHO SOFTWARE... 100</b>	
	<b>ZÁVĚR .....</b>	<b>103</b>
	<b>CONCLUSION .....</b>	<b>105</b>
	<b>SEZNAM POUŽITÉ LITERATURY.....</b>	<b>107</b>
	<b>SEZNAM OBRÁZKŮ .....</b>	<b>111</b>
	<b>SEZNAM TABULEK.....</b>	<b>114</b>
	<b>SEZNAM PŘÍLOH.....</b>	<b>115</b>



## SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK

## ZKRATKY Z OBLASTI ŘÍZENÍ KVALITY

SPC	Statistické řízení procesů
SQC (SŘJ)	Statistické řízení kvality (jakosti)
QDA	System automatizovaného sběru dat
MSA	Analýza měřicího systému
CL	Střední průměr
UCL	Horní regulační mez
LCL	Dolní regulační mez
UWL	Horní výstražná mez
LWL	Dolní výstražná mez
USL	Horní specifikační mez
LSL	Dolní specifikační mez
ČSN	Česká státní norma
CPI	Index způsobilosti procesu
ISO	Mezinárodní organizace pro standardizaci
QMS	System managementu jakosti
EWMA	Metoda exponenciálně vážených klouzavých průměrů
CUSUM	Metoda kumulovaných součtů
ARIMA	Autoregresní integrovaný proces klouzavých průměrů
SARIMA	Sezónní ARIMA
SACC	Výběrový autokorelační diagram
SCC	Shewhartův regulační diagram
MR	Klouzavé rozpětí
S-W test	Shapiro-Wilkův test
NC	Přípustné procento zmetků
Ppm	NC (nevyhovujících) výrobků vzhledem k milionu kusů

## MATEMATICKÉ FUNKCE A VÝRAZY

$\mathbb{N}$ ( $\mathbb{R}$ ) [ $\mathbb{Z}$ ]	Množina přirozených (reálných) [komplexních] čísel
$\mathbf{A}_k$	Čtvercová matice typu $k \times k$
$ \mathbf{A} $	Determinant matice $\mathbf{A}$
$X$	Náhodná veličina $X$

$F(X)$	Distribuční funkce náhodné veličiny $X$
$E(X)$	Střední hodnota náhodné veličiny $X$
$\text{var}(X)$	Rozptyl náhodné veličiny $X$
$\text{cov}(X, Y)$	Kovariance náhodných veličin $X$ a $Y$
$\text{corr}(X, Y)$	Korelační koeficient náhodných veličin $X$ a $Y$
$\bar{X}$	Aritmetický průměr náhodné veličiny $X$
$H_0$	Nulová hypotéza
$H_A, H_1$	Alternativní hypotéza
$X \sim N(0, 1)$	Náhodná veličina $X$ má asymptoticky normální rozdělení se střední hodnotou 0 a rozptylem 1
$\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$	Proces $\varepsilon_t$ je proces bílého šumu se střední hodnotou 0 a rozptylem $\sigma_\varepsilon^2$
$\chi_s^2(\alpha)$	$(1 - \alpha)$ kvantil $\chi^2$ rozdělení s $s$ stupni volnosti
$\frac{\partial f(x, y)}{\partial x}$	Parciální derivace funkce $f(x, y)$ podle proměnné $x$
$t_s(\alpha)$	$(1 - \alpha)$ kvantil Studentova $t$ rozdělení s $s$ stupni volnosti
$\xrightarrow{d}$	Konvergence v distribuci
$\xrightarrow{p}$	Konvergence v pravděpodobnosti
$x \wedge y$	$\min(x, y)$
$x \vee y$	$\max(x, y)$
$\phi(z)$	Distribuční funkce normovaného normálního rozdělení
$[x]$	Celočíselná část $x$
$\langle X, X \rangle_t, [X, X]_t$	Kvadratická variace procesu přidružená k $X_t$
$V_t(X)$	Jednoduchá variace procesu $X$
$f_{v_i}(v_1, v_2, \dots, v_n)$	$\frac{\partial}{\partial v_i} f(v_1, v_2, \dots, v_n)$
$f_{v_i, v_j}(v_1, v_2, \dots, v_n)$	$\frac{\partial^2}{\partial v_i \partial v_j} f(v_1, v_2, \dots, v_n)$ , atd.
$\partial_\theta f(v_1, v_2, \dots, v_n; \theta)$	$\frac{\partial}{\partial \theta} f(v_1, v_2, \dots, v_n; \theta)$
$\partial_\theta^k f(v_1, v_2, \dots, v_n; \theta)$	$\frac{\partial^k}{\partial \theta^k} f(v_1, v_2, \dots, v_n; \theta)$
$\prod_n(A)$	Rozdělení intervalu $A = [a, b]$ na $n$ subintervalů $[a = x_0, x_1),$ $[x_1, x_2), \dots, [x_{n-1}, x_n = b]$
$\ \Pi_n\ $	$\max_j  x_{j+1} - x_j $

## ÚVOD

Klasické metody statistického řízení procesů (SPC), jako jsou Shewhartovy regulační diagramy a CUSUM regulační diagramy předpokládají, že získaná procesní data jsou nezávislá. Nicméně, tento předpoklad byl zpochybněn, jakmile bylo zjištěno, že v mnoha praktických situacích, jsou data sériově korelována. Výkonnost klasických regulačních diagramů se v případě autokorelace v datech výrazně zhorší. Statistické metody v řízení procesů pro monitorování sériově korelovaných nebo autokorelovaných procesů získaly velkou pozornost v literatuře, týkající se statistického řízení kvality. Zaměření většiny výzkumných studií je však na detekci posunu střední hodnoty procesu. Detekce změny ve variabilitě a autokorelační struktuře časové řady v důsledku některých zvláštních příčin variability ovlivňující systém je často přehlížena.

V úvodní kapitole praktické části této práce nejprve představím systém automatizovaného sběru dat v podniku Continental Barum, s.r.o., kde nyní působím jako statistik. Výstupem takovýchto systémů je v mnoha případech časová řada autokorelovaných pozorování, což má za následek špatnou implementaci klasických Shewhartových regulačních diagramů. Následkem této skutečnosti jsou z těchto dat vypočteny zkreslené hodnoty indexů způsobilosti a výkonnosti procesu, jako jsou například  $C_p$ ,  $C_{pk}$ ,  $P_{pk}$  a  $C_{pm}$ .

Další nosnou kapitolou praktické části budou tvořit ukázky monitorování autokorelovaných procesů na praktických případových studiích z reálných procesů ze své osobní praxe v oblasti SPC.

Hlavním cílem této diplomové práce je srovnání jednak aparátu, navrženého Atienzou a Tangem (2002) a aparátu navrženého Montgomerym a Friedmanem (1989) s klasickým Shewhartovým regulačním diagramem po stránce výkonnosti v simultánním monitorování střední hodnoty, rozptylu a autokorelační struktury sériově korelovaných procesů. Dále se zaměřím se na studium chování regulačního diagramu  $\lambda_{LS,t}$  navrženého Atienzou et al. (1998) používaného pro detekci úrovně posunu v procesu AR(1). Výkonnost tohoto aparátu bude vyhodnocena a srovnána s Alvanovým a Robertsovým aparátem SCC (diagram zvláštních příčin – modifikovaný Shewhartův regulační diagram pro rezidua). Pro srovnání výkonnosti těchto regulačních diagramů použiji simulaci Monte Carlo implementovanou v programovém prostředí jazyka R, která se nachází v knihovně `spc`.

Pro zpracování této diplomové práce jsem použil programy Statgraphics Centurion, Minitab 16, NCSS, MS Excel 2007 a jazyk R.

## **I. TEORETICKÁ ČÁST**

## 1 INTERPRETACE REGULAČNÍHO DIAGRAMU

Pro interpretaci regulačního diagramu platí obecně základní pravidla:

- a) Leží-li všechny body uvnitř UCL a LCL, je proces pokládán za statisticky zvládnutý a není vyžadován žádný zásah do procesu.
- b) Leží-li některý bod mimo regulační mez UCL nebo LCL, je proces pokládán za statisticky nezvládnutý, je vyžadována identifikace vymezené příčiny této odchylky a přijetí opatření s cílem úplné či alespoň částečné eliminace vymezeného vlivu.

Použijí-li se i meze výstražné, mohou nastat kromě uvedené základní situace ještě další dvě situace:

- 1) Některý bod leží uvnitř výstražných mezí – lze předpokládat, že proces je ve statisticky zvládnutém stavu a není třeba žádného zásahu.
- 2) Některý bod leží mezi UWL a UCL, resp. mezi LWL a LCL. V této situaci se doporučuje postupovat následovně: Ihned bez ohledu na kontrolní interval se provede další výběr. Jestliže nový bod, odpovídající tomuto bezprostřednímu výběru, leží mezi výstražnými mezemi, není třeba do procesu zasahovat. Jestliže však i tento nový bod leží mimo výstražné meze, je to signál, že na proces s velkou pravděpodobností působí vymezená příčina a je nutné provést regulační zásah. Nejčastěji používané testy nenáhodných seskupení jsou uvedeny v příloze P I. (Tošenovský a Noskiewičová, 2000)

### 1.1 Obecný postup sestavení a analýzy regulačního diagramu

Dosavadní poznatky o regulačních diagramech můžeme stručně shrnout do devíti základních kroků, které je nutné provádět bez ohledu na použitou metodu SPC. Jsou to tyto kroky:

1. Volba regulované veličiny.
2. Sběr a záznam dat.
3. Ověření předpokladů o datech.
4. Volba rozsahu výběru.
5. Volba vhodného regulačního diagramu.

6. Výpočet hodnot zvoleného testového kritéria (výběrové charakteristiky) pro jednotlivé výběry.
7. Ověření a zajištění statistické zvládnutosti procesu.
8. Ověření a zabezpečení způsobilosti procesu.
9. Vlastní regulace procesu. (Tošenovský a Noskiewičová, 2000)

## 1.2 Klasické Shewhartovy regulační diagramy

Autorem je americký odborník W. A. Shewhart (1924). Jde o grafickou pomůcku umožňující oddělit identifikovatelné (zvláštní) příčiny od náhodných (obecných) příčin variability procesu. Konstrukce regulačních diagramů má matematicko-statistický základ. Klasické regulační diagramy patří do skupiny regulačních diagramů bez paměti, neboť v aktuální hodnotě použitého testového kritéria nezohledňují jeho předchozí hodnoty. Proto se tyto diagramy hodí zejména pro odhalování větších sporadických odchylek v procesu (odchylek větších než  $2\sigma$  od požadované úrovně). (Tošenovský a Noskiewičová, 2000)

Shewhartovy regulační diagramy patří do skupiny regulačních diagramů tzv. „bez paměti“, neboť při analýze aktuální hodnoty regulované veličiny neberou v úvahu její předchozí hodnoty. Shewhartovy regulační diagramy se dělí na dva základní typy. Regulační diagramy pro regulaci měření, které se používají pro měřitelný znak kvality resp. technologický parametr procesu, pro nějž budeme často používat pojem regulovaná veličina. Pro správnou aplikaci těchto diagramů se požaduje, aby regulovaná veličina, která je spojitou náhodnou veličinou, měla normální rozdělení s konstantní střední hodnotou a konstantním rozptylem. Dále se předpokládá, že jednotlivá měření jsou vzájemně nezávislá, což značí, že naměřená hodnota nezávisí na předchozích naměřených hodnotách.

Statistickou stabilitu procesu charakterizují parametry  $\mu_0$  a  $\sigma_0$  regulované veličiny. Parametr  $\mu_0$ , tj. střední hodnota, představuje hodnotu regulované veličiny, na niž je výrobní proces skutečně nastaven, parametr  $\sigma_0$ , tj. směrodatná odchylka, charakterizuje přesnost výrobního procesu. Tyto předpoklady je nutné před volbou metody regulace měřením ověřit, a pokud je některý z nich porušen, je třeba zvolit jiný typ regulačního diagramu.

Jsou-li uvedené předpoklady splněny, pak se volí takové regulační diagramy, v nichž se simultánně sledují jak stabilita polohy, tak stabilita stejnoměrnosti rozptylu regulované veličiny. Pro tyto účely zde uvedeme regulační diagramy, označené  $(\bar{x}, R)$ ,  $(\bar{x}, s)$  a  $(x_i, R_{kl,i})$ .

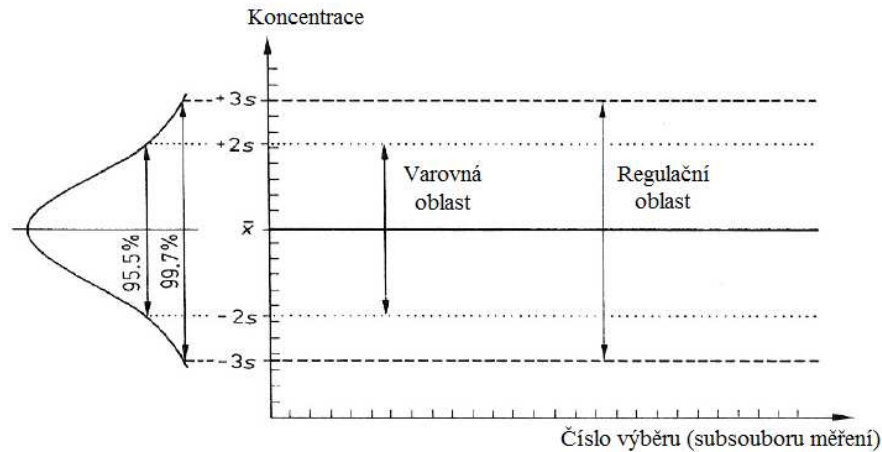
Chceme-li sledovat pouze stabilitu polohy procesu, použijeme některý z regulačních diagramů uvedených na prvním místě ve dvojici, zajímá-li nás pouze míra stejnoměrnosti rozptylu procesu, vybereme některý z regulačních diagramů uvedených na druhém místě. Regulační diagramy pro regulaci srovnáváním. Ty se používají tehdy, sledují-li se počty neshodných produktů resp. počty neshod na těchto produktech, tedy regulovaná veličina je diskrétní náhodnou veličinou. Tyto regulační diagramy se dělí na dvě skupiny.

– Jestliže se zjišťuje počet neshod na jednotlivých produktech, pak se používá buď regulační diagram pro počet neshod, označený  $c$  nebo regulační diagram pro počet neshod na jednotku, označený  $u$ . Příkladem použití těchto diagramů může být to, že ve sklárně se vyrábí stejné tabule skel, na nichž se kontroluje počet vad (např. bublin).

– Zjišťují-li se počty neshodných produktů ve výběru, pak se používá buď regulační diagram pro počet neshodných produktů ve výběru, označený  $np$  nebo regulační diagram pro podíl neshodných produktů ve výběru, označený  $p$ . Příkladem použití těchto diagramů může být automatická výrobní linka, na níž se vyrábí série určitých produktů. Při kontrole se z této série vybírají kontrolní skupiny, přičemž se určuje počet neshodných produktů mezi těmito vybranými. (Tošenovský a Noskiewičová, 2000)

*Základní postup při konstrukci Shewhartova regulačního diagramu je následující:*

1. Zvolíme takovou část procesu, která odpovídá naší představě, předpisu, nebo zkušenosti a připravíme příslušná procesní data.
2. Na základě těchto dat stanovíme jejich statistický model představovaný odhadem střední hodnoty (aritmetickým průměrem) a směrodatnou odchylkou a ověříme platnost statistických předpokladů Shewhartova diagramu.

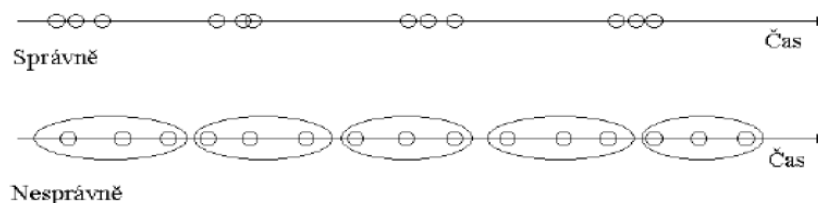


**Obr. 1.1** Shewhartův diagram (Zdroj: Vlastní zpracování)

3. Z těchto dvou parametrů se zkonstruuje vlastní regulační diagram, který má podobu základní linie CL a horní a spodní regulační meze LCL a UCL.
4. Do tohoto regulačního diagramu se pak vynášejí data z procesu a sleduje se výskyt „zvláštních případů“ signalizujících nečekanou změnu chování procesu, z nichž základní je překročení regulační meze.
5. Výskyt zvláštních případů (viz příloha P II) se eviduje a hledá se tzv. přiřaditelná příčina, která je způsobila, ke každému výskytu zvláštního případu by se měla zapsat přiřaditelná příčina (pokud se jí podaří identifikovat) a opatření, které bylo přijato. (ČSN ISO 7870, 1995)

### "Racionální podskupiny"

*Podskupina* = opakovaná měření v "jednom" časovém okamžiku, volba podskupin výrazně ovlivňuje správnou funkci regulačního diagramu. Časový rozsah hodnot podskupiny má být malý ve srovnání s intervalem mezi podskupinami. Podskupina však musí odrážet variabilitu, jinak se regulační meze příliš zúží a regulační diagram je nepoužitelný.



**Obr. 1.2** Výběr racionální podskupiny (Zdroj: Tošenovský a Noskiewičová, 2000)



### 1.3 Účinnost regulačního diagramu

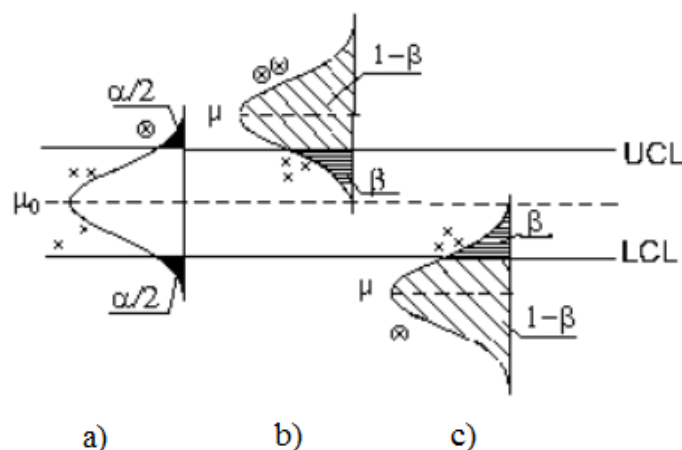
Jak již bylo zmíněno, klasické regulační diagramy patří do skupiny regulačních diagramů bez paměti, neboť v aktuální hodnotě použitého testové kritéria nezohledňují jeho předchozí hodnoty. Proto se tyto diagramy hodí zejména pro odhalování větších sporadických odchylek v procesu (odchylek větších než  $2\sigma$  od požadované úrovně). Stanovení regulačních mezí je při konstrukci regulačního diagramu klíčovým problémem ovlivňujícím účinnost regulačního diagramu. Proto se budeme této problematice při výkladu klasických Shewhartových diagramů věnovat podrobněji. (Tošenovský a Noskiewičová, 2000)

#### Účinnost regulačního diagramu lze vyjádřit následovně:

Regulační meze jsou stanoveny v takové vzdálenosti od střední přímký, aby:

- signál o tom, že proces je statisticky nestabilní (nevládnutý), i když ve skutečnosti je statisticky stabilní, byl vydáván co nejméně často (co nejmenší riziko zbytečného signálu  $\alpha$ ),
- signál o tom, že došlo k nepřijatelné odchylce procesu, byl vydán co nejdříve (co nejmenší riziko chybějícího signálu  $\beta$ , resp. co největší pravděpodobnost odhalení nepřijatelné odchylky  $1 - \beta$ ).

Význam rizik  $\alpha$  a  $\beta$ , resp. pravděpodobnosti  $1 - \beta$  je patrný z obrázku 1.3.



$\mu_0$  je požadovaná průměrná úroveň regulované veličiny

$\mu_1$  je průměrná úroveň regulované veličiny po změně procesu

**Obr. 1.3** Riziko zbytečného signálu  $\alpha$  a riziko chybějícího signálu  $\beta$

(Zdroj: Tošenovský a Noskiewičová, 2000)

Hodnota  $\alpha$  se nazývá *riziko zbytečného signálu* a představuje pravděpodobnost zbytečného hledání vymezeného vlivu na základě informace z regulačního diagramu o tom, že proces není ve statisticky zvládnutém stavu (např. bod mimo akční meze), i když ve skutečnosti k žádné významné změně procesu nedošlo (obr. 1.3a). S tímto nesprávným závěrem jsou spojeny náklady na pokus najít příčinu neexistujícího problému. Hodnota  $\beta$  se nazývá *riziko chybějícího signálu* a je to pravděpodobnost, že regulační diagram neodhalí nepřijatelnou změnu procesu včas (všechny body daného testového kritéria leží uvnitř regulačních mezí a netvoří žádné nenáhodné seskupení). S tímto nesprávným závěrem jsou zase spojeny náklady vyvolané tím, že se do procesu včas nezasáhlo. Na obr. 1.3b a 1.3c je situace zobrazena pro posun střední úrovně regulované veličiny z  $\mu_0$  na  $\mu_1$ . Hodnota  $(1 - \beta)$ , obecně nazývaná *silofunkce* testu, zde vyjadřuje pravděpodobnost detekce nepřijatelné změny procesu, která má být včas odhalena. Za předpokladu, že neexistuje žádná chyba měření, lze riziko  $\alpha$  (tj. jeho pravou hodnotu), stanovit ze vztahu (1.1).

Jak již bylo řečeno, akční regulační meze vymezují pásmo působení pouze náhodných příčin. Jsou tedy specifikovány tak, že testové kritérium (výběrová charakteristika  $f(x_j)$ ) bude při statisticky zvládnutém procesu ležet mimo regulační mez se zvolenou malou pravděpodobností  $\alpha$  (tuto pravděpodobnost jsme definovali jako riziko zbytečného signálu). Při oboustranné regulaci pak platí:

$$P\left[f(x_j) \notin \langle LCL, UCL \rangle | SZP\right] = \alpha, \quad (1.1)$$

$SZP$  je proces ve statisticky zvládnutém stavu.

Pokud má testové kritérium symetrické rozdělení, pak LCL a UCL leží ve stejné vzdálenosti od CL a vztah (1.1) se rozpadne na dvě samostatné rovnice:

$$P\left[f(x_j) > UCL | SZP\right] = \alpha/2 \quad (1.2)$$

$$P\left[f(x_j) < LCL | SZP\right] = \alpha/2 \quad (1.3)$$

Při jednostranné regulaci platí:

$$P\left[f(x_j) > UCL | SZP\right] = \alpha \quad (1.4)$$

Nebo

$$P\left[f(x_j) < LCL | SZP\right] = \alpha. \quad (1.5)$$

Se zmenšováním vzdálenosti mezi regulačními mezemi při zachování rozsahu výběru  $n$  riziko  $\alpha$  stoupá. To může vést ke snaze mít pásmo mezi regulačními mezemi co nejširší. Výsledkem však může být diagram nedostatečně citlivý na významné odchylky v procesu s vysokým rizikem chybějícího signálu  $\beta$  (riziko  $\beta$  roste se zvětšující se vzdáleností mezi regulačními mezemi). Obecně je třeba přijmout kompromis, který bere v potaz oba druhy rizik. Klasické Shewhartovy diagramy uvedené v normě ČSN ISO 8258 však zohledňují pouze riziko  $\alpha = 0,0027$  (to znamená, že přibližně ve 3 případech z 1000 bude bod v diagramu ležet mimo některou akční mez, i když v procesu nedošlo k žádné změně). Při této hodnotě rizika  $\alpha$  jsou akční regulační meze UCL a LCL ve vzdálenosti  $\pm 3$  směrodatné odchylky daného testového kritéria od střední přímky CL. (Tošenovský a Noskievičová, 2000)

Chceme-li zvýšit citlivost klasického Shewhartova regulačního diagramu na menší odchylky, než umožňují  $\pm 3\sigma$  akční meze, je nutné při stanovení akčních regulačních mezí vzít v potaz i riziko  $\beta$ . Cílem bude opět toto riziko minimalizovat. Riziko chybějícího signálu  $\beta$  závisí na velikosti procesu, kterou chceme co nejrychleji odhalit. Stanovení rizika  $\beta$  si ukážeme na příkladu změny střední hodnoty měřitelné regulované veličiny z požadované úrovně  $\mu_0$  na nepřijatelnou úroveň  $\mu_s$ . Použitým testovým kritériem je výběrový průměr, který má asymptoticky normální rozdělení, proto pro akční regulační meze platí vztahy:

$$UCL = \mu_0 + 3\sigma / \sqrt{n}, \quad (1.6a)$$

$$LCL = \mu_0 - 3\sigma / \sqrt{n}, \quad (1.6b)$$

$\sigma$  je směrodatná odchylka regulované veličiny (často neznámá),

$n$  je rozsah výběru,

$\sigma / \sqrt{n}$  je směrodatná odchylka výběrového průměru.

Před změnou střední hodnoty má rozdělení výběrových průměrů parametru  $\mu_0$  a  $(\sigma^2 / n)$ ,

po změně má rozdělení testového kritéria parametry  $\mu_s$  a  $(\sigma^2 / n)$ .

Po  $\mu_s > \mu_0$  (obr. 1.3b) vyjdeme při určení  $\beta$  ze vztahu:

$$\beta = P(\bar{X} \leq UCL) = F(UCL) = \Phi\left(\frac{UCL - \mu_s}{\sigma/\sqrt{n}}\right), \quad (1.6c)$$

$F(UCL)$  je hodnota distribuční funkce normálního rozdělení v bodě  $x = UCL$ ,

$\Phi\left(\frac{UCL - \mu_s}{\sigma/\sqrt{n}}\right)$  je hodnota distribuční funkce normovaného normálního rozdělení v bodě  $(UCL - \mu_s)/(\sigma/\sqrt{n})$ . (Tošenovský a Noskiewičová, 2000)

V praxi se častěji hodnota rizika  $\beta$  volí rozumně nízká a hledá se rozsah výběru  $n$ , při kterém je zvolená úroveň  $\beta$  zaručená. Vhodný rozsah výběru  $n$  pro  $\mu_s > \mu_0$  lze určit ze vztahu:

$$n = \left(\frac{\sigma(u_{1-\beta} + 3)}{\mu_s - \mu_0}\right)^2, \quad (1.6d)$$

kde  $u_{1-\beta}$  je  $100 \cdot (1-\beta)\%$  kvantil normovaného normálního rozdělení odpovídající zvolené úrovni rizika  $\beta$ . (Tošenovský a Noskiewičová, 2000)

Podrobnější informace o vztazích mezi riziky  $\alpha$  a  $\beta$  a způsobech zvyšování citlivosti regulačního diagramu lze nalézt v (Wright, 1995).

Co se týká posunu střední hodnoty, můžeme riziko  $\beta$  určit dle následujícího vztahu:

$$\beta = P\{LCL \leq \bar{x} \leq UCL \mid \mu = \mu_1 = \mu_0 = k\sigma\}. \quad (1.7)$$

Vzhledem k tomu, že  $\bar{x} \sim N(\mu, \sigma^2/n)$  a spodní a horní regulační meze jsou  $UCL = \mu_0 + L\sigma/\sqrt{n}$  a  $LCL = \mu_0 - L\sigma/\sqrt{n}$ , můžeme vztah pro riziko  $\beta$  rozepsat následovně:

$$\begin{aligned} \beta &= \Phi\left[\frac{UCL - (\mu_0 + k\sigma)}{\sigma/\sqrt{n}}\right] - \Phi\left[\frac{LCL - (\mu_0 + k\sigma)}{\sigma/\sqrt{n}}\right] = \\ &= \Phi\left[\frac{\mu_0 + L\sigma/\sqrt{n} - (\mu_0 + k\sigma)}{\sigma/\sqrt{n}}\right] - \Phi\left[\frac{\mu_0 - L\sigma/\sqrt{n} - (\mu_0 + k\sigma)}{\sigma/\sqrt{n}}\right], \end{aligned} \quad (1.8)$$

kde  $\Phi$  značí distribuční funkci normovaného normálního rozdělení. Tento vztah se po drobnějších úpravách zredukuje na tvar

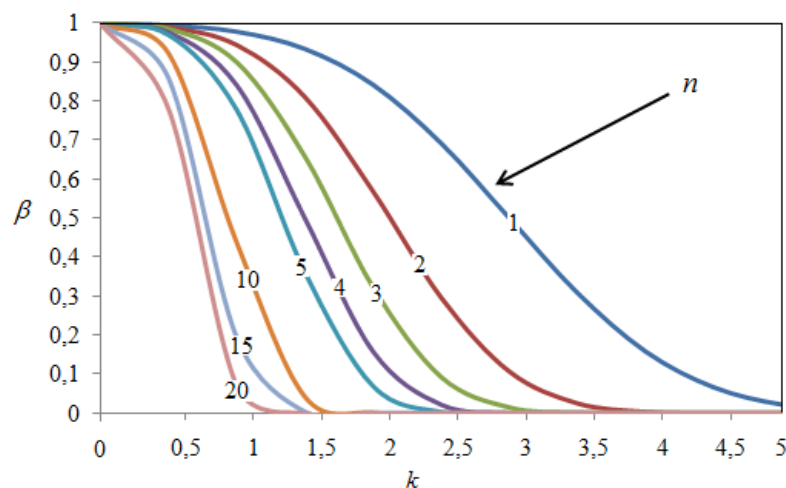
$$\beta = \Phi(L - k\sqrt{n}) - \Phi(-L - k\sqrt{n}). \quad (1.9)$$

Pro ilustraci užití předchozího vztahu budeme předpokládat, že použijeme regulační diagram pro  $\bar{x}$  s  $L = 3$  (obvyklé rozpětí mezi  $3\sigma$ ), rozsah výběru  $n = 5$  a chceme určit pravděpodobnost odhalení posunu na  $\mu_1 = \mu_0 + 2\sigma$  u prvního výběru. Potom, za předpokladu, že  $L = 3$ ,  $k = 2$  a  $n = 5$ , dostaneme (Tošenovský a Noskiewiczová, 2000)

$$\beta = \Phi[3 - 2\sqrt{5}] - \Phi[-3 - 2\sqrt{5}] = \Phi(-1,147) - \Phi(-7,37) \cong 0,0708.$$

Hodnota  $\beta$  je pravděpodobnost, že regulační diagram neodhalí nepřijatelnou změnu procesu včas a hodnota  $(1 - \beta)$ , která vyjadřuje pravděpodobnost detekce nepřijatelné změny procesu, která má být včas odhalena je potom  $1 - \beta = 1 - 0,0708 = 0,9292$ . Křivka operační charakteristiky pro různé posuny a velikosti výběrů je pro případ rozsahu mezi  $3\sigma$  ( $L = 3$ ) znázorněna na následujícím obrázku.

Tento obrázek ukazuje, že pro rozsahy 4, 5 a 6, diagram  $\bar{x}$  není nijak zvlášť efektivní v detekci malých posunů, řekněme do rozsahu  $1,5\sigma$  a méně, po posunu v prvním výběru. Například, pokud máme posun  $1,0\sigma$  a  $n = 5$ , potom z následujícího obrázku dostáváme přibližně  $\beta = 0,75$ . Takže pravděpodobnost, že posun bude odhalen v prvním výběru je pouze  $1 - \beta = 0,25$ . (Montgomery, 2009)



**Obr. 1.4** Křivky operační charakteristiky pro regulační diagram  $\bar{x}$  s mezemi v rozpětí  $3\sigma$ .  $\beta = P$  (pravděpodobnost, že nebude zjištěn posun střední hodnoty o  $k\sigma$  včas) (Zdroj: Vlastní zpracování)

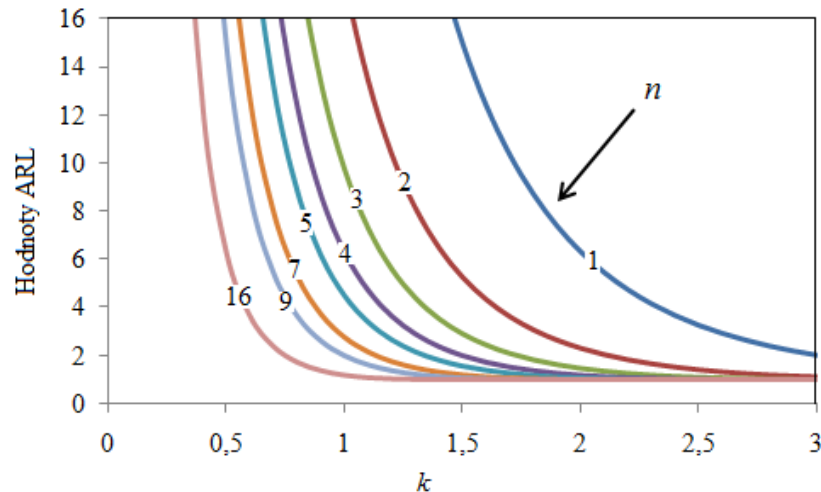
Nicméně, pravděpodobnost odhalení posunu ve druhém výběru je  $\beta(1 - \beta) = 0,75(0,25) = 0,19$ , kdežto pravděpodobnost, že je odhalena ve třetím výběru je

$\beta^2(1-\beta) = (0,75^2)0,25 = 0,14$ . Takže pravděpodobnost, že posun bude odhalen v následujícím  $r$  tém výběru se stanoví jednoduše  $\beta^{r-1}(1-\beta)$ . Riziko  $\alpha$  a  $\beta$  se u některých SPC technik uplatňují nepřímo pomocí hodnoty ARL (Average Run Length – průměrný počet výběrů vedoucí k signálu). Je-li proces statisticky zvládnutý, pak ARL se značí  $ARL(0)$ . Mezi  $ARL(0)$  a rizikem zbytečného signálu  $\alpha$  platí vztah:  $ARL(0) = 1/\alpha$ , protože ke zbytečnému signálu dojde v průměru po  $1/\alpha$  výběrech. ARL je střední hodnota rozdělení náhodné proměnné RL (Run Length – počet výběrů vedoucí k signálu). (Montgomery, 2009)

Průměrný počet výběrů mezi okamžikem vzniku odchylky v procesu o normované velikosti  $\delta$  a okamžikem jejího odhalení v regulačním diagramu se značí  $ARL(\delta)$ . Mezi  $ARL(\delta)$  a rizikem chybějícího signálu  $\beta$  je vztah:  $ARL = \sum_{r=1}^{\infty} r\beta^{r-1}(1-\beta) = \frac{1}{1-\beta}$ , neboť k signálu, že došlo k odchylce o normované velikosti  $\delta$ , dojde v průměru po  $1/(1-\beta)$  výběrech od okamžiku vzniku této odchylky. Tento jednoduchý vztah platí pouze pro klasické Shewhartovy regulační diagramy. Stanovení hodnot  $ARL(\delta)$  např. u metod CUSUM nebo EWMA vyžaduje řešení pomocí složitějšího matematického aparátu.

Pro náš příklad tedy platí  $ARL(\delta) = 1/(1-\beta) = 1/0,25 = 4$ . Jinými slovy, očekávaný počet výběrů vedoucí k signálu posunu  $1\sigma$  při  $n = 5$  je čtyři. (Montgomery, 2009)

Na následujícím obrázku jsou zobrazeny křivky ARL pro rozsahy výběrů  $n = 1, 2, 3, 4, 5, 7, 9$  a  $16$  pro regulační diagram  $\bar{x}$ , kde ARL je, z hlediska očekávaného počtu výběrů, přijata za účelem zjištění posunu. Pro ilustraci slouží následující obrázek, všimněte si, že pokud chceme zjistit posun o  $1,5\sigma$  pomocí rozsahu výběru  $n = 3$ , potom průměrný počet potřebných výběrů bude  $ARL_1 = 3$ . Všimněte si také, že bychom mohli zredukovat  $ARL_1$  na hodnotu přibližně 1, pokud bychom zvýšili rozsah výběru na  $n = 16$ . (Montgomery, 2009)



**Obr. 1.5** Hodnoty ARL pro diagram  $\bar{x}$  s mezemi o rozpětí  $3\sigma$  v případech posunu procesu o  $k\sigma$  (Zdroj: Vlastní zpracování)

Při navrhování regulačního diagramu je snaha volit  $ARL(0)$  co největší, aby ztráty spojené se zbytečným vyhledáváním příčin neexistujících odchylek nebo se zbytečným seřizováním procesu byly minimální. Naopak  $ARL(\delta)$  se volí co nejmenší, aby se minimalizovaly ztráty spojené s pozdním odhalením odchylky. (Montgomery, 2009)

#### 1.4 Členění regulačních diagramů

Metody SPC dnes představují širokou škálu prostředků analýzy procesu. V tabulce, která je uvedena v příloze P III, je členění regulačních diagramů dle vybraných hledisek. (Fabian et al., 2007)

Regulační diagramy lze členit podle různých hledisek, některá zde budou uvedena. Podle charakteru regulované veličiny dělíme regulační diagramy na:

- *regulační diagramy pro regulaci měřením*, používané v případě, kdy je regulovaná veličina spojitou náhodnou veličinou,
- *regulační diagramy pro regulaci srovnáváním*, používané v případě, kdy je regulovaná veličina diskrétní náhodnou veličinou.

Podle počtu znaků jakosti simultánně sledovaných na jedné jednotce výběru dělíme regulační diagramy na:

- regulační diagramy pro sledování jednoho znaku jakosti (klasické Shewhartovy diagramy),

- regulační diagramy pro sledování více znaků jakosti najednou (např. Hotellingův diagram, MEWMA, MCUSUM).

Jestliže při výpočtu aktuální hodnoty regulované veličiny se neuvažují její předchozí získané hodnoty, jsou to regulační diagramy bez paměti (např. klasické Shewhartovy diagramy), pokud se předchozí hodnoty regulované veličiny uvažují, jde o regulační diagramy s paměti (např. diagramy CUSUM, EWMA). (Fabian et al., 2007)

## 1.5 Shewhartovy regulační diagramy pro regulaci měřením

Většina postupů analýzy dat a z ní plynoucí závěry jsou závislé na splnění základních předpokladů, za nichž byly tyto postupy odvozeny. Nejsou-li základní předpoklady splněny, jsou veškeré další klasické postupy jako výpočet průměru, intervalů spolehlivosti, kvantilů, většiny testů, klasických Shewhartových regulačních diagramů, konstrukce modelů pro popis časové řady, atd. zpochybnitelné a napadnutelné. (Meloun a Militký, 2004)

Poskytují obvykle nesprávné výsledky a závěry. Typická porušení předpokladů pro aplikaci regulace Shewhartovými diagramy nebo v různých technologiích, ale také i pro konstrukci modelů pro popis časových řad, je uvedeno v následující tabulce.

Zmíněné předpoklady je nutné ověřit pomocí statistických testů a jedná se o tyto *základní předpoklady pro statistickou regulaci procesu*:

- normalita rozdělení dat, symetrie,
- konstantní střední hodnota procesu,
- konstantní rozptyl (směrodatná odchylka) dat,
- nezávislost, nekorelovanost dat,
- nepřítomnost vybočujících hodnot. (Meloun a Militký, 2004)

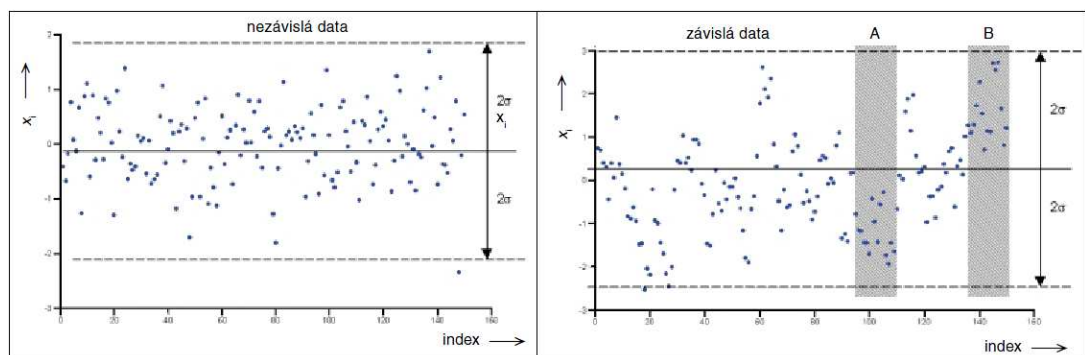
Přehled typických porušení předpokladů pro aplikaci regulace Shewhartovými diagramy v různých technologiích je uveden v příloze P I.



## 1.6 Vliv autokorelace na klasické regulační diagramy

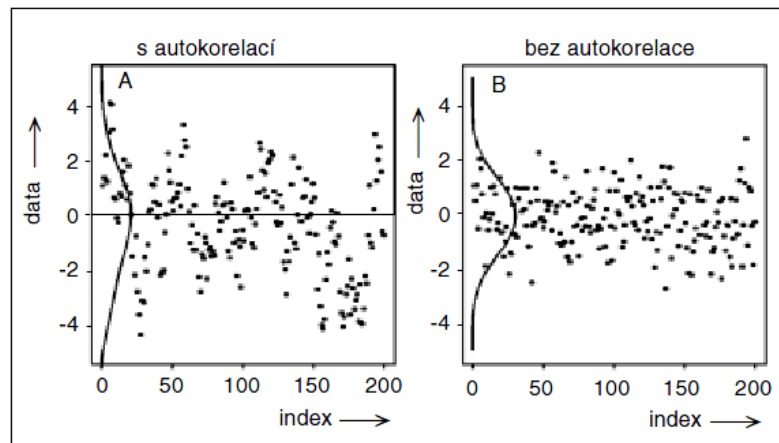
Hlavním předpokladem efektivní aplikace klasických Shewhartových regulačních diagramů je vzájemná nezávislost hodnot sledovaného znaku kvality. I velmi nízký stupeň vzájemné závislosti hodnot sledovaného znaku kvality (autokorelace) vyvolává selhání klasických Shewhartových regulačních diagramů. Selhání má podobu vysokého počtu falešných signálů. Tento jev není vůbec výjimečnou záležitostí v případě spojitých procesů, kde je autokorelace vyvolána velkou setrvačností procesu v čase. Stále častějším fenoménem se však vzájemná závislost dat stává i v podmínkách diskrétních procesů. Důvody lze spatřovat v automatizaci výrobních, zkušebních a kontrolních postupů, což umožňuje získat data z každého produktu (a je-li to potřebné, nejen jednoho znaku kvality), tedy nejen z výběru  $n$  produktů odebraných z procesu po uplynutí určité doby  $T_v$  (kontrolního intervalu) od předchozího výběru, jak je obvyklé při realizaci sběru a záznamu dat při aplikaci klasických Shewhartových diagramů a jejich modifikací. Existuje několik postupů, které vedou k odstranění autokorelace dat a umožňují využití SPC i v podmínkách vzájemné závislosti hodnot sledovaného znaku kvality. Čtyři možnosti řešení uvedeného problému jsou diskutovány v (Noskievičová, ©2008). Jedná se o tyto metody:

1. metoda prodloužení kontrolního intervalu;
2. postup s využitím aparátu modelování časových řad pomocí ARIMA modelování;
3. aproximační postup založený na využití statistiky EWMA;
4. dynamický EWMA diagram.



**Obr. 1.6** Závislost v datech způsobuje vzrůst zdánlivé variability a nepoužitelnost Shewhartova diagramu (Zdroj: Noskievičová, ©2008)

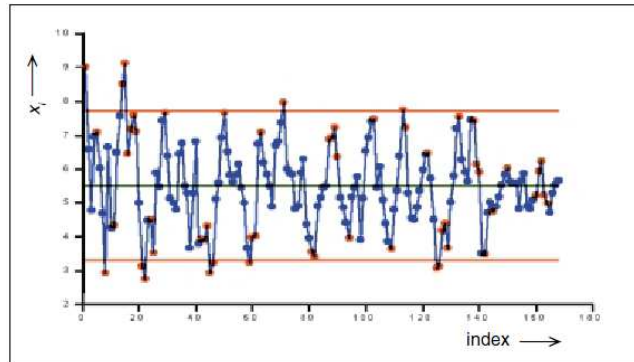
V případě závislých dat lze sice numericky vyčíslit směrodatnou odchylku podle známého vztahu, avšak pro tuto odchylku již nebude platit, že v intervalu o šířce  $4s$  ( $\pm 2s$ ) kolem střední hodnoty se očekává výskyt hodnoty s pravděpodobností asi 95 %. Ve vyšrafované oblasti A na obr. 1.6 vpravo se výskyt hodnot v horní polovině intervalu vůbec neočekává, v oblasti B se zase nečeká výskyt hodnot v dolní polovině intervalu. Horní mez intervalu v oblasti A a spodní mez intervalu v oblasti B postrádají smysl. Následkem toho zde nelze mimo jiné použít žádný z klasických (a často nesmyslně vyžadovaných) Shewhartových diagramů, popř. diagramů CUSUM (Kupka, ©2001). Potenciál zmenšení variability (a tím i zvětšení tzv. indexu způsobilosti  $C_p$  a kvality vůbec) lze vyjádřit v případě, že proces lze popsat autokorelačním modelem 1. řádu, podílem



**Obr. 1.7** Autokorelace v datech může mít přiřaditelnou příčinu, po jejímž odstranění se výrazně zmenší variabilita (Zdroj: Noskiewičová, ©2008)

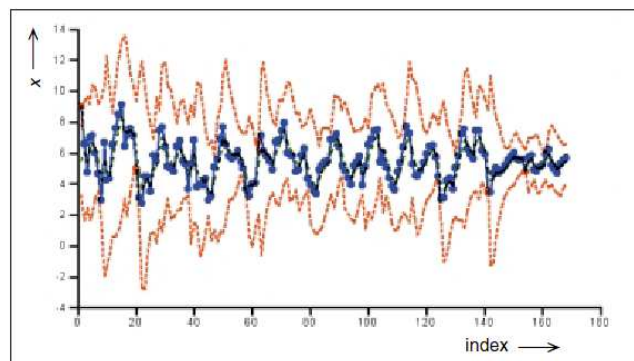
$$\frac{s}{s_0} = \sqrt{1 - r^2}, \quad (1.10)$$

kde  $s_0$  je směrodatná odchylka vypočítaná podle klasického vztahu,  $s$  je minimální směrodatná odchylka, které by bylo docíleno využitím znalosti modelu procesu (tedy znalostí  $r$ ) a  $r$  je odhad autokorelačního koeficientu 1. řádu (Noskiewičová, ©2008). V případě závislých dat na obr. 1.6 a 1.7, kde  $r = 0,8$ , by bylo  $s_0/s = 0,6$ , což odpovídá nárůstu  $C_p$  o 40 % a zmenšení Taguchiho ztráty o 64 %. Nepoužitelnost Shewhartova diagramu  $x$ -individual pro autokorelovaná data ilustruje obr. 1.8. (Noskiewičová, ©2008)



**Obr. 1.8** Shewhartův diagram  $x$ -individual vykazuje nesmyslných 53 porušení pravidel v důsledku přítomnosti autokorelace (Zdroj: Noskiewičová, ©2008)

Příčinou této skutečnosti je výpočet regulačních mezí ze směrodatné odchylky procesu získané z klouzavých rozpětí (viz ISO ČSN 8225 - Shewhartovy regulační diagramy). Taková směrodatná odchylka však odpovídá spíše směrodatné odchylce  $s$  z předchozího vzorce, kdežto skutečné hodnoty  $x$  odpovídají směrodatné odchylce  $s_0$ . Toto je jedna z nejčastějších příčin nespokojenosti se Shewhartovými diagramy. V takovém případě je nutné použít alternativní regulační diagram, např. dynamický diagram EWMA (viz obr. 1.9), aplikovaný na téže data, diagram AR nebo diagram ARMA, které berou závislost dat v úvahu. (Noskiewičová, ©2008)



**Obr. 1.9** Diagram EWMA, popř. diagram AR, reálně diagnostikuje pět pravidel porušení (Zdroj: Noskiewičová, ©2008)

Je třeba zdůraznit, že tento problém existuje i v případě diagramu  $x$ -průměr, kde se hodnotí chování průměrů racionálních podskupin (několika měření) místo jednotlivých hodnot. Zde je kromě autokorelace problém ještě častá párová korelace mezi jednotlivými měřeními v podskupině. V takovém případě Shewhartův diagram opět selhává a je vhodné použít vícerozměrný Hotellingův diagram, založený na kovarianční matici (místo prosté směrodatné odchylky) a Mahalanobisově vzdálenosti. (Kupka, ©2001)

## 2 REGULAČNÍ DIAGRAMY PRO ČASOVÉ ŘADY

Tato kapitola bude zaměřena na popis matematického aparátu, na kterém jsou založeny regulační diagramy pro časové řady. Budou zde popsány základní principy těchto diagramů, které budou aplikačně demonstrovány na praktických případových studiích z reálných procesů v praktické části této diplomové práce.

### 2.1 Metoda kumulovaných součtů (CUSUM)

Vymežitelné příčiny variability procesu lze rozdělit na základní typy, a to na:

1. Příčiny sporadické
2. Příčiny přetrvávající.

Příčiny sporadické vznikají náhle, působí na proces pouze krátkou dobu a pak se ztrácejí, i když se opět mohou vyskytnout. Odchytky způsobené těmito příčinami bývají větší a efektivně mohou být indikovány klasickými Shewhartovými regulačními diagramy.

Příčiny přetrvávající vyvolávají náhlé, určitou dobu trvající odchytky v parametrech rozdělení regulované veličiny. Pokud odchytky nejsou většího rozsahu, nejsou klasické Shewhartovy regulační diagramy schopny rychle a hospodárně (tj. při malých rozsazích výběrů) tyto odchytky odhalit. Je to dáno tím, že přetrvávající příčiny nemohou být odhaleny pomocí jednotlivých izolovaných výběrů. Jejich rychlé a hospodárné odhalení vyžaduje metodu, která akumuluje informaci obsaženou ve všech za sebou jdoucích výběrech a poskytuje tak jasnější obraz o změně procesu. Kromě toho je problematické pomocí klasického Shewhartova regulačního diagramu odhadnout s dostatečnou mírou přesností okamžik vzniku odchytky a její velikosti. (Tošenovský a Noskiewičová, 2000)

Tyto požadavky splňuje metoda kumulovaných součtů (Cumulative Sum Control Charts), zkráceně CUSUM. Tyto regulační diagramy vytvořil Page v roce 1954. Výhody diagramů CUSUM lze ve srovnání s klasickými Shewhartovými regulačními diagramy formulovat následovně:

- Diagramy CUSUM jsou citlivější na změny procesu malé a střední velikosti ( $0,5\sigma - 2,0\sigma$ ).
- Metoda CUSUM je hospodárnější pro riziko zbytečného signálu  $\alpha \leq 0,1$  a tato hospodárnost se projevuje tím více, čím je  $\alpha$  menší.

- Metoda CUSUM indikuje uvedené malé a střední změny procesu 2 – 4x rychleji při stejném rozsahu výběru  $n$ .
- Metoda CUSUM je spojena s nižšími náklady na kontrolu (menšími rozsahy výběru  $n$ ) při stejném riziku  $\alpha$ .
- Metoda CUSUM umožňuje přesněji určit okamžik počátku změny parametrů rozdělení regulované veličiny, odhadnout její velikost a určit směr působení.

Nevýhodou metody CUSUM je vyšší pracnost konstrukce diagramu a jeho vyhodnocení. Výhody uvedené metody se nejvýrazněji uplatní ve fázi vlastní statistické regulace, i když může účinně pomoci i při rozboru chování procesu před zavedením regulace a k retrospektivní analýze. (Tošenovský a Noskievičová, 2000)

### 2.1.1 Základní charakteristiky metody CUSUM

Při aplikaci metody CUSUM se konstruuje diagram, do něhož se na ose  $x$  zaznamenává pořadí výběru  $k$  a na ose  $y$  se vynášejí hodnoty testového kritéria  $Y_k$ . Hodnotu této výběrové charakteristiky po  $k$  tém výběru lze stanovit následovně:

$$y_k = -k \cdot Y_0 + \sum_{j=1}^k f(x_j) = \sum_{j=1}^k [f(x_j) - Y_0] = y_{k-1} + [f(x_k) - Y_0], \quad y_0 = 0, \quad (2.1)$$

$Y_0$  je cílová (žádaná) hodnota sledovaného parametru rozdělení regulované veličiny,

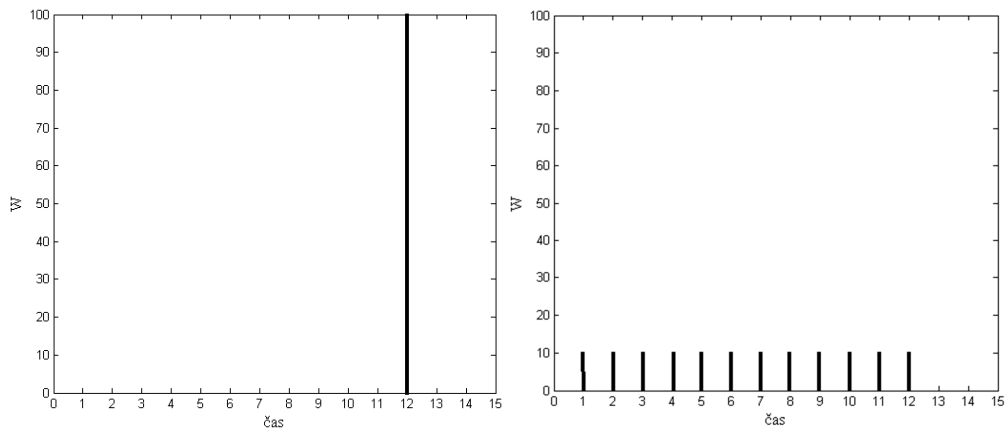
$x_j$  je vektor  $n$  hodnot sledované regulované veličiny (sledovaného znaku kvality) zjištěných v  $j$  tém výběru ( $j = 1, 2, \dots, k$ ),

$f(x_j)$  je funkcí  $n$  hodnot regulované veličiny v  $j$  tém výběru, tj. hodnota zvolené výběrové charakteristiky v  $j$  tém výběru,

$k$  je pořadí výběru.

Pokud  $Y_0 = \mu_0$  ( $\mu_0$  je požadovaná úroveň střední hodnoty regulované veličiny) a funkce  $f(x_j)$  je výběrový průměr  $\bar{x}_j$ , pak jde o diagram CUSUM pro výběrové průměry. Ve speciálním případě  $n = 1$  jde o diagram CUSUM pro individuální hodnoty. Kromě uvedených diagramů existují též diagramy pro výběrové rozpětí, počet vadných prvků, resp. počet vad na definovaném objektu. (Tošenovský a Noskievičová, 2000)

Podíváme-li se na výraz (2.1), můžeme testové kritérium  $Y_k$  interpretovat jako kumulativní součet odchylek zvolené výběrové charakteristiky od cílové hodnoty pro všechny dosud provedené výběry. To řadí metodu CUSUM do skupiny diagramů s pamětí. Vlastnosti paměti diagramu jsou patrné z obrázku 2.1, kde v části vlevo jsou zobrazeny vlastnosti paměti klasického Shewhartova diagram a v části vpravo pak diagramu CUSUM. (Tošenovský a Noskiewičová, 2000)



**Obr. 2.1** Srovnání diagramů z hlediska vlastností jejich paměti zobrazených pomocí průběhu váhových koeficientů  $w$  (%)  
(Zdroj: Tošenovský a Noskiewičová, 2000)

Vlastnosti paměti diagramů jsou vyjádřeny pomocí průběhu váhových koeficientů  $w$ . Shewhartův diagram bez aplikace testů nenáhodných seskupení pracuje pouze s poslední hodnotou použité výběrové charakteristiky (např. v klasickém Shewhartově diagramu pro výběrové průměry se o tom, zda proces je či není statisticky zvládnutý, rozhoduje pouze na základě hodnoty výběrového průměru vypočteného z  $n$  hodnot regulované veličiny v dosud posledním, aktuálním výběru). To znamená, že poslední hodnotě je přiřazena váha  $w = 100\%$ . Klasické Shewhartovy regulační diagramy jsou tedy diagramy bez paměti. U diagramu CUSUM se všechny dosud vypočtené hodnoty odchylek použité výběrové charakteristiky (např. výběrového průměru) od cílové hodnoty promítají do poslední (aktuální) hodnoty testového kritéria  $y_k$  se stejnou váhou. To znamená, že paměť diagramu CUSUM je neomezená a rovnoměrná. Podrobná charakteristika metody CUSUM bude dále uvedena pro diagram CUSUM pro výběrové průměry. (Tošenovský a Noskiewičová, 2000)

### 2.1.2 Definování diagramu CUSUM pro výběrové průměry ( $\bar{x}$ )

Jak již bylo uvedeno, v tomto diagramu se cílová hodnota  $Y_0$ , rovná požadované úrovni střední hodnoty rozdělení regulované veličiny  $\mu_0$ . V této souvislosti je třeba upozornit na skutečnost, že pokud se zvolená hodnota  $\mu_0$  výrazně liší od skutečného středu procesu, vede to k nevhodnému průběhu bodů v diagramu CUSUM (body mohou opustit diagram) a posouzení změn procesu je pak znesnadněno či dokonce nemožné. Dále se předpokládá normální rozdělení regulované veličiny s konstantním rozptylem.

Hodnotu testového kritéria po  $k$  tém výběru  $y_k$ , pak můžeme určit ze vztahu:

$$C_k = \sum_{j=1}^k (\bar{x}_j - \mu_0) = C_{k-1} + (\bar{x}_k - \mu_0), \quad C_0 = 0, \quad (2.2)$$

kde  $k$  je pořadí výběru ( $k = 1, 2, \dots$ ),

$\bar{x}_j$  je výběrový průměr z hodnot regulované veličiny v  $j$  tém výběru ( $j = 1, 2, \dots, k$ ).

I když další výklad předpokládá oboustrannou regulaci při použití diagramu CUSUM pro výběrové průměry, lze obdobně interpretovat i ostatní digramy CUSUM. (Tošenovský a Noskiewičová, 2000)

### 2.1.3 Interpretace průběhu diagramu CUSUM

Interpretace průběhu diagramu CUSUM vychází z následující úvahy:

- Jestliže proces je udržován na cílové hodnotě  $\mu_0$ , pak body v diagramu zachovávají směr přibližně rovnoběžný s osou  $x$ .
- Jestliže došlo k náhlé změně střední hodnoty regulované veličiny přibližně v době, kdy byl odebrán  $q$ -tý výběr, a tato změna přetrvává, pak body v diagramu počínaje bodem  $[q, C_q]$  náhodně oscilují kolem přímky, která není rovnoběžná s osou  $x$ .

Jestliže střední hodnota procesu roste nebo klesá a ještě se nestabilizovala (v procesu existuje trend), pak body v diagramu tvoří křivku viditelně se zakřívující nahoru nebo dolů. (Tošenovský a Noskiewičová, 2000)

## 2.2 Diagramy EWMA

Obdobně jako diagram CUSUM uvedený v předchozí kapitole, také diagram EWMA se hodí pro situace, kdy v procesu dochází k náhlým malým, ale přetrvávajícím změnám procesu a hodnoty sledovaného znaku jakosti nejsou závislé. Na rozdíl od klasických diagramů závisí regulační meze na okamžiku výběru. Autorem je Roberts (1959). Diagram EWMA pracuje s testovým kritériem  $Y_k$ , jehož hodnota po  $k$  tém výběru je definována následovně:

$$y_k = (1 - \lambda)^k \cdot Y_0 + \lambda \cdot \sum_{j=1}^k (1 - \lambda)^{k-j} \cdot f(x_j) \quad \text{pro } j = 1, 2, \dots, k \text{ a pro } 0 < \lambda < 1, \quad (2.3)$$

$f(x_j)$  je hodnota zvolené výběrové charakteristiky v  $j$  tém výběru,

$k$  je pořadí výběru,

$Y_0$  je požadovaná úroveň parametru rozdělení regulované veličiny.

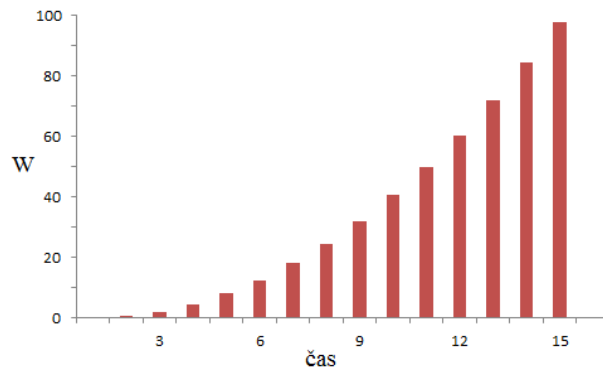
Jestliže platí  $Y_0 = \mu_0$  ( $\mu_0$  je požadovaná úroveň střední hodnoty regulované veličiny) a funkce  $f(x_j)$  je výběrový průměr  $\bar{x}_j$ , pak jde o diagram EWMA pro výběrové průměry. V případě, že rozsah výběru  $n = 1$ , pak jde o diagram EWMA pro individuální hodnoty. Obdobně jako u většiny ostatních metod statistické regulace procesu rozlišujeme diagramy EWMA pro jednostrannou a oboustrannou regulaci. (Tošenovský a Noskievičová, 2000)

Vztah (2.3) lze upravit na zjednodušený tvar:

$$y_k = (1 - \lambda) \cdot y_{k-1} + \lambda \cdot f(x_k) \quad \text{pro } 0 < \lambda < 1, y_0 = Y_0. \quad (2.4)$$

Diagramy EWMA patří k diagramům s neomezenou nerovnoměrnou pamětí. Vlastnosti paměti určuje parametr  $\lambda$  ( $0 < \lambda < 1$ ). Je-li  $\lambda = 1$ , pak je testované kritérium EWMA identické s testovaným kritériem v klasickém Shewhartově diagramu. To znamená, že váha je dána pouze poslední „nejčerstvější“ hodnotě použité výběrové charakteristiky (obr. 2.2). Naopak, čím více se  $\lambda$  blíží k nule, tím více se blíží vlastnosti paměti diagramu EWMA vlastnostem paměti diagramů CUSUM. Zobrazení vlastnosti diagramů EWMA je na následujícím obrázku. (Tošenovský a Noskievičová, 2000)





**Obr. 2.2** Vlastnosti paměti diagram EWMA zobrazené pomocí průběhu váhových koeficientů  $w$  (%) (Zdroj: Tošenovský a Noskiewičová, 2000)

Podrobnější analýza diagramů EWMA bude dále provedena na diagramu EWMA pro výběrové průměry ( $\bar{x}$ ) při oboustranné regulaci. (Tošenovský a Noskiewičová, 2000)

### 2.2.1 Klasický diagram EWMA pro výběrové průměry ( $\bar{x}$ )

Klasický diagram EWMA se obdobně jako diagram CUSUM hodí pro situaci, kdy v procesu dochází k náhlým malým, ale přetrvávajícím změnám procesu a hodnoty sledovaného znaku jakosti nejsou závislé. Klasický diagram EWMA je definován rozsahem výběru a parametru  $\lambda$  a  $k$ . Stejně jako v případě diagramu CUSUM zohledňuje diagram EWMA jak riziko zbytečného signálu  $\alpha$ , tak riziko chybnějšího signálu  $\beta$ .

V klasickém digramu EWMA pro výběrové průměry ( $\bar{x}$ ) se hodnota  $Y_0$  rovná požadované úrovni střední hodnoty regulované veličiny  $\mu_0$ . Předpokládá se normální rozdělení regulované veličiny s konstantním rozptylem. Hodnota testovaného kritéria po  $k$  tém výběru pak stanovíme následovně:

$$EWMA_k = (1 - \lambda)^k \cdot \mu_0 + \lambda \cdot \sum_{j=1}^k (1 - \lambda)^{k-j} \cdot \bar{x}_j \quad \text{pro } 0 < \lambda < 1, j = 1, 2, \dots, k, \quad (2.5)$$

resp.

$$EWMA_k = (1 - \lambda)EWMA_{k-1} + \lambda \cdot \bar{x}_k, \quad EWMA_0 = \mu_0 \quad \text{pro } 0 < \lambda < 1. \quad (2.6)$$

Volbou parametru  $\lambda$  je ovlivňována míra zapomínání. Čím  $\lambda$  zvolíme větší, tím menší je vliv starších výběrových průměrů. Čím více se blíží hodnota parametru  $\lambda$  k 1, tím více se diagram EWMA blíží klasickému Shewhartovu regulačnímu diagramu, který nemá žádnou paměť. Čím více se blíží  $\lambda$  k nule (nikdy se nerovná nule), tím více se váhové faktory

EWMA diagramu blíží váhovým faktorům v diagramu CUSUM. (Tošenovský a Noskievičová, 2000)

Střední přímka CL se v EWMA diagramu pro výběrové průměry se stanoví ze vztahu:

$$CL = \mu_0 \quad (2.7)$$

Regulační meze se určí následovně:

$$UCL = CL + K \cdot \sigma_{EWMA} = \mu_0 + K \cdot \sigma_{EWMA} \quad (2.8)$$

$$LCL = CL - K \cdot \sigma_{EWMA} = \mu_0 - K \cdot \sigma_{EWMA} \quad (2.9)$$

Směrodatná odchylka  $\sigma_{EWMA}$  se stanoví dle vztahu:

$$\sigma_{EWMA} = \frac{\sigma_0}{\sqrt{n}} \sqrt{\frac{\lambda}{2-\lambda} \cdot [1 - (1-\lambda)^{2k}]}, \quad (2.10)$$

kde  $\sigma_0$  je požadovaná úroveň směrodatné odchylky regulované veličiny,

$K$  je konstanta pro stanovení regulačních mezí vyhovujících zvolenému riziku  $\alpha$ , resp. zvolené hodnotě ARL(0). (Tošenovský a Noskievičová, 2000)

Na rozdíl od klasických Shewhartových regulačních diagramů regulační meze v diagramu EWMA závisí na časovém okamžiku výběru, avšak relativně rychle přecházejí do asymptotického tvaru:

$$UCL_a = CL + K \cdot \sigma_a = \mu_0 + K \cdot \sigma_a \quad (2.11)$$

$$LCL_a = CL - K \cdot \sigma_a = \mu_0 - K \cdot \sigma_a. \quad (2.12)$$

Směrodatná odchylka  $\sigma_a$  se stanoví ve vztahu:

$$\sigma_a = \frac{\sigma_0}{\sqrt{n}} \sqrt{\frac{\lambda}{2-\lambda}}. \quad (2.13)$$

### Postup sestavení a analýza klasického diagramu EWMA

Při konstrukci a analýze, klasického regulačního diagramu EWMA lze postupovat dle jednoduché metodiky, kterou navrhnul Crowder, a která využívá pro stanovení parametrů  $\lambda$  a  $K$  nomogramů. Kroky uvedené metodiky budou demonstrovány v následující případové studii. (Tošenovský a Noskievičová, 2000)

## 2.3 ARIMA A SARIMA procesy

### 2.3.1 Oblast použití

Automatizované procesy s vlastním systémem automatizovaného monitorování procesu včetně sběru a zpracování dat přinášejí nové podmínky, které znemožňují použití klasických Shewhartových regulačních diagramů např.:

- sledovaný znak kvality se zjišťuje na každé vyrobené jednotce a výrobní cykly jsou velmi krátké (tyto skutečnosti vedou k autokorelaci dat),
- náklady na měření a kontrolu jsou malé,
- na každé vyrobené jednotce je často sledován více než jeden znak kvality.

Pro uvedení situace jsou vhodné regulační diagramy ARIMA nebo dynamický diagram EWMA. (Noskiewičová, ©2008)

Stochastický proces je v čase uspořádaná řada náhodných veličin  $\{y(s, t), s \in S, t \in T\}$ , kde  $S$  je výběrový prostor a  $T$  je indexní řada. Pro každé  $t \in T$  je  $y(\cdot, t)$  náhodná veličina definovaná na výběrovém prostoru  $S$ . Pro každé  $s \in S$  je  $y(s, \cdot)$  realizace stochastického procesu definovaná na indexní řadě  $T$ , tj. uspořádaná řada čísel, z nichž každé odpovídá jedné hodnotě indexní řady. Časovou řadu lze tedy chápat jako realizaci stochastického procesu. V dalších částech budeme předpokládat, že indexní řada je řadou celých čísel, tj.  $T = \{0, \pm 1, \pm 2, \dots\}$ . Pro zjednodušení budeme stochastické procesy značit jako  $\{y_t, t = 0, \pm 1, \pm 2, \dots\}$  resp.  $\{y_t\}$  a časové řady jako  $y_t$ . Stochastický proces je striktně stacionární, jestliže je jeho pravděpodobnostní chování invariantní vůči posunům v čase. Protože striktní stacionaritu je v praxi obtížné ověřovat, byl v analýze časových řad zaveden méně omezující pojem slabá stacionarita stochastických procesů. V následujícím textu budou definovány procesy ARIMA a SARIMA. (Arlt et al., ©2002)

### 2.3.2 Procesy ARIMA

Model ARIMA, se používá tehdy, vykazuje-li po transformaci integrovaného procesu pomocí difference  $d$  tého řádu výsledný proces takové autokorelace a parciální autokorelace, že jej lze vyjádřit ve formě stacionárního a invertibilního modelu ARMA( $p, q$ ). Potom se původní integrovaný proces vyjádřený ve formě

$$\phi_p(B)(1-B)^d y_t = \theta_q(B) a_t \quad (2.14)$$

nazývá autoregresní integrovaný proces klouzavých průměrů řádu  $p$ ,  $d$ ,  $q$  a označuje se jako  $ARIMA(p,d,q)$ . Vlastnosti tohoto typu procesu jsou obdobné jako vlastnosti „náhodné procházky“ (Random Walk). Tyto procesy se někdy nazývají pouze integrovanými procesy řádu  $d$  a označují se jako  $I(d)$ . (Arlt et al., ©2002)

### 2.3.3 Sezónní procesy – Procesy SARIMA

Myšlenka sezónního procesu je následující: jako v případě procesu  $ARIMA$  předpokládáme vzájemnou závislost mezi veličinami  $\dots y_{t-3}, y_{t-2}, y_{t-1}, y_t, y_{t+1}, y_{t+2}, y_{t+3}, \dots$ , a protože tento proces obsahuje ještě sezónní kolísání, lze očekávat i závislost mezi sobě odpovídajícími veličinami v jednotlivých sezónách, tj. mezi veličinami  $\dots y_{t-2s}, y_{t-1s}, y_{ts}, y_{t+1s}, y_{t+2s}, \dots$ , kde  $s$  je délka sezónní periody (např. u měsíčních časových řad 12, u čtvrtletních 4). Budeme vycházet z předpokladu, že proces obsahuje oba typy závislostí. Závislost uvnitř period je zachycena modelem  $ARIMA$

$$\phi_p(B)(1-B)^d y_t = \theta_q(B) b_t. \quad (2.15)$$

Proces  $\{b_t\}$ , pak obsahuje pouze sezónní závislosti a může být popsán modelem

$$\Phi_p(B^s)(1-B^s)^D b_t = \Theta_Q(B^s) a_t, \quad (2.16)$$

kde  $\Phi_p(B^s) = 1 - \Phi_1 B^s - \dots - \Phi_p B^{ps}$  a  $\Theta_Q(B^s) = 1 - \Theta_1 B^s - \dots - \Theta_Q B^{Qs}$ . Prostřednictvím členu  $(1 - B^s)$  se konstruují sezónní diference. Jestliže se procesy (2.15) a (2.16) spojí, získá se proces

$$\Phi_p(B^s) \phi_p(B)(1-B)^d (1-B^s)^D y_t = \theta_q(B) \Theta_Q(B^s) a_t, \quad (2.17)$$

který je označován jako  $SARIMA(p,d,q)(P,D,Q)_s$ , kde  $p$  je řád procesu  $AR$ ,  $q$  řád procesu  $MA$ ,  $d$  řád prosté diference,  $P$  řád sezónního procesu  $AR$ ,  $Q$  řád sezónního procesu  $MA$ ,  $D$  řád sezónní diference a  $s$  je délka sezónní periody. Podmínky stacionarity a invertibility u sezónní části jsou obdobné jako u části nesezónní. (Arlt et al., ©2002)

V případě čtvrtletních časových řad lze sezónní integrovaný proces prvního řádu vyjádřit ve tvaru

$$(1-B^4)y_t = a_t. \quad (2.18)$$

Polynomiální rovnice

$$(1-B^4) = (1-B)(1+B)(1+iB)(1-iB) = 0 \quad (2.19)$$

má čtyři kořeny: 1, -1, i, -i. Tyto kořeny leží na jednotkové kružnici ve frekvencích 0,  $\pi$ ,  $\pi/2$  a  $3\pi/2$ . Frekvence  $\pi/2$  a  $3\pi/2$  jsou ve čtvrtletních časových řadách nerozlišitelné, proto se uvažuje pouze frekvence  $\pi/2$ . Frekvence 0 je nesezónní frekvencí a jednotkový kořen v této frekvenci znamená přítomnost stochastického trendu, frekvence  $\pi$  znamená, že každý rok obsahuje dva cykly a frekvence  $\pi/2$ , že každý rok obsahuje jeden cyklus. Z uvedeného vyplývá, že sezónní diferencování neznamená pouze odstraňování tzv. integrované sezónní složky, ale také odstraňování stochastického trendu, neboť sezónně integrovaný proces (2.17) zahrnuje rovněž nesezónní integrovaný proces typu náhodné procházky. (Arlt et al., ©2002)

### 2.3.4 Identifikace a ověřování modelu, konstrukce předpovědi

#### 2.3.4.1 Identifikace modelu

Jednou z nejtěžších úloh při výstavbě Boxových-Jenkinsových modelů je jejich identifikace. Tato úloha spočívá v rozhodnutí, jaký typ modelu vybrat. Jde o nelehkou činnost, která je v mnoha případech závislá na citu a zkušenosti analytika. Identifikace je přitom pouze první fází konstrukce modelů, neboť identifikovaný model je třeba ještě ověřit a upravit. Podívejme se nyní na jednotlivé kroky identifikace modelu.

Uvažujme model ARIMA(p,d,q) ve formě (2.15).

1. Nejprve je vhodné prozkoumat graf časové řady. V mnoha případech je možné na první pohled rozpoznat přítomnost trendu. V této fázi jde především o subjektivní zhodnocení situace. Nicméně, již na základě tohoto zhodnocení je možné stacionarizovat časovou řadu či provést jiné úpravy, jako je linearizace časové řady, resp. její stabilizace z hlediska rozptylu pomocí logaritmické transformace. Tento typ transformace je vhodné provádět před vlastním diferencováním časové řady. Důvodem je skutečnost, že diferencováním je možné získat i záporné hodnoty.

2. Dalším krokem je výpočet odhadů ACF a PACF původní časové řady. Na jejich základě je možné potvrdit, že časovou řadu je třeba stacionarizovat (v případě, že hodnoty výběrové ACF a PACF v prvním zpoždění jsou velmi blízké jedné a ostatní hodnoty výběrové ACF klesají pomalu).
3. Po stacionarizaci časové řady se použijí výběrové ACF a PACF pro identifikaci modelů AR a MA (nalezení hodnot  $p$  a  $q$ ). Tato identifikace je založena na principu podobnosti výběrových ACF a PACF s teoretickými ACF a PACF. Následující tabulka obsahuje popis tvarů ACF a PACF pro modely AR, MA a ARMA. (Arlt et al., ©2002)

**Tab. 7.8** Tvary ACF a PACF modelů AR, MA a ARMA (Zdroj: Arlt et al., ©2002)

Model	ACF	PACF
AR( $p$ )	Exponenciální a / nebo exponenciálně sinusoidní pokles	$\varphi_{kk} = 0$ pro $k > p$
MA( $q$ )	$\rho_k = 0$ pro $k > q$	Omezená exponenciálním a / nebo exponenciálně sinusoidním poklesem
ARMA( $p, q$ )	Od zpoždění ( $q - p$ ) pro $q > p$ exponenciální nebo exponenciálně sinusoidní pokles	Od zpoždění ( $p - q$ ) pro $p > q$ omezená exponenciálním nebo exponenciálně sinusoidním poklesem

Obsahuje-li časová řada také sezónní složku, je třeba identifikovat model typu SARIMA tvaru (2.17). Princip je stejný jako v předchozím případě. (Arlt et al., ©2002)

1. Nejprve je třeba časovou řadu stacionarizovat. Pokud je to nutné, linearizuje se řada pomocí logaritmické transformace. Poté se řada diferencuje, je třeba přitom mít na paměti, že sezónní diference zahrnuje rovněž diferenci prostou. Stacionarizaci pomocí sezónní diference indikuje tvar výběrové ACF a PACF. Tyto funkce jsou charakteristické vysokými hodnotami v nesezónních a sezónních frekvencích.
2. V další fázi se vypočítá výběrová ACF a PACF pro stacionarizovanou časovou řadu, s jejich pomocí se potom určí typ sezónních modelů SAR, SMA, či SARMA (v sezónních frekvencích mají tyto funkce statisticky významně odlišné hodnoty od

nuly, tyto hodnoty však nejsou tak vysoké, aby bylo možné časovou řadu považovat za nestacionární). Po identifikaci a určení modelu sezónní složky je třeba vypočítat výběrovou ACF a PACF tentokrát pro rezidua sezónního modelu a na jejich základě posoudit, zda by bylo vhodné model doplnit o složku AR, MA či ARMA. (Arlt et al., ©2002)

### 2.3.4.2 Diagnostická kontrola modelu

#### Testování nesystematické složky

V modelu ARIMA( $p, d, q$ ) typu (2.15) předpokládáme, že  $\{a_t\}$  je proces bílého šumu. Základní empirická diagnostická kontrola spočívá v posouzení reziduí

$$\hat{a}_t = \left[ \hat{\theta}_q(B) \right]^{-1} \phi_p(B) y_t, \quad (2.20)$$

kde  $\hat{\phi}_p(B) = 1 - \hat{\phi}_1 B - \dots - \hat{\phi}_p B^p$  a  $\hat{\theta}_q(B) = 1 - \hat{\theta}_1 B - \dots - \hat{\theta}_q B^q$ .

Autokorelaci nesystematické složky lze testovat pomocí výběrové autokorelační funkce

$$r_k = \frac{\sum_t \hat{a}_t \hat{a}_{t-k}}{\sum_t \hat{a}_t^2}. \quad (2.21)$$

Není-li nesystematická složka autokorelovaná, měly by hodnoty této funkce ležet uvnitř intervalu  $\pm 2 / \sqrt{T}$  (95% interval spolehlivosti). (Arlt et al., ©2002)

Další možností jak zjistit, zda nesystematická složka není autokorelovaná, je použití *Portmanteau* testu, který navrhli Box a Pierce. Testuje se hypotéza  $H_0: \rho_1 = \rho_2 = \dots = \rho_K = 0$  proti hypotéze  $H_1: \text{non } H_0$ , kde  $\rho_k$ ,  $k = 1, \dots, K$  jsou autokorelace nesystematické složky modelu pro zpoždění  $k$ . Je-li model ARIMA správně zkonstruovaný, potom má statistika (Arlt et al., ©2002)

$$Q = T \sum_{k=1}^K \hat{r}_k^2, \quad (2.22)$$

pro vysoké  $T$  a  $K$  přibližně rozdělení  $\chi^2$  s  $(K - p - q)$  stupni volnosti. (Arlt et al., ©2002)

### Testování parametrů modelu

Nevykazuje-li nesystematická složka autokorelaci, je vhodné ještě otestovat jednotlivé parametry v modelu, tj. parametry  $\mu, \phi_i, i = 1, 2, \dots, p$  a  $\theta_i, i = 1, 2, \dots, q$ . Testování se provádí na základě  $t$  testů, resp. pomocí statistik

$$t_{\mu_i} = \frac{\hat{\mu}_i}{\hat{S}_{\mu}}, t_{\phi_i} = \frac{\hat{\phi}_i}{\hat{S}_{\phi_i}}, i = 1, 2, \dots, p \text{ a } t_{\theta_i} = \frac{\hat{\theta}_i}{\hat{S}_{\theta_i}}, i = 1, 2, \dots, q, \quad (2.23)$$

kde  $\hat{S}_{\mu}, \hat{S}_{\phi_i}, \hat{S}_{\theta_i}$  jsou odhady směrodatných chyb odhadů jednotlivých parametrů modelu ARIMA. (Arlt et al., ©2002)

#### 2.3.4.3 Výpočet předpovědi

Při výpočtu předpovědi budeme vycházet ze skutečnosti, že předpověď  $y_T(h)$  je dána podmíněnou střední hodnotou  $E(y_{T+h} | y_T, y_{T-1}, y_{T-2}, \dots)$ . Předpokládejme model ARMA( $p, q$ ) ve formě

$$y_t = \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}. \quad (2.24)$$

Pro  $t = T + h$  dostaneme

$$y_{T+h} = \phi_1 y_{T+h-1} + \dots + \phi_p y_{T+h-p} + a_{T+h} - \theta_1 a_{T+h-1} - \dots - \theta_q a_{T+h-q}. \quad (2.25)$$

Budeme-li uvažovat podmíněné střední hodnoty v čase  $T$

$$y_T(h) = \phi_1 y_T(h-1) + \dots + \phi_p y_T(h-p) + a_T(h) - \theta_1 a_T(h-1) - \dots - \theta_q a_T(h-q), \quad (2.26)$$

kde

$$y_T(j) = E(y_{T+j} | y_T, y_{T-1}, \dots), j \geq 1,$$

$$y_T(j) = y_{T+j}, j \leq 0,$$

$$a_T(j) = 0, j \geq 1,$$

$$a_T(j) = y_{T+j} - y_{T+j-1}(1) = a_{T+j}, j \leq 0. \text{ (Arlt et al., ©2002)}$$



## 2.4 Regulační diagram ARIMA

Klasická Shewhartova koncepce SPC předpokládá, že naměřená data nejsou autokorelovaná. I velmi nízký stupeň autokorelace dat způsobuje selhání klasických Shewhartových regulačních diagramů. Selhání má podobu velkého počtu bodů mimo regulační meze v regulačním diagramu. Tento jev není výjimečný v případě spojitých procesů, kde je autokorelace dat dána velkou setrvačností procesů v čase (chemické procesy, klimatologické procesy apod.). Stále častějším fenoménem se však autokorelace dat stává i v podmínkách diskrétních procesů, a to zejména výrobních procesů s krátkými výrobními cykly, velkou rychlostí výroby a s vysokým stupněm automatizace výrobních, zkušebních a kontrolních operací. V těchto podmínkách lze získat data o každém produktu s tím důsledkem, že doba mezi měřeními (záznamem) dvou po sobě jdoucích hodnot sledované veličiny je velmi krátká. Jednou z možností, jak řešit otázku autokorelovaných dat, je stochastická koncepce modelování časových řad při použití autoregresních integrovaných modelů klouzavých průměrů, tzv. modelů ARIMA. V rámci stochastické koncepce tvorby lineárních autoregresních modelů (*autoregressive*, tzv. modely AR), modelů klouzavých průměrů (*mean average*, modely MA), smíšených modelů (modely ARMA) a modelů ARIMA, vycházející z Boxovy-Jenkinsovy metodiky, je časová řada chápána jako realizace stochastického procesu. Boxova-Jenkinsova metodika tak představuje moderní koncepci analýzy stacionárních a nestacionárních časových řad, založenou na teorii pravděpodobnosti. Lineární modely AR, MA a ARMA jsou nástrojem modelování stacionárních procesů. Tyto modely mají charakteristický tvar autokorelační funkce (*Autocorrelation Function* – ACF) a parciální autokorelační funkce (*Partial Autocorrelation Function* – PACF), které jsou základním nástrojem poskytujícím informace o stochastickém procesu. Odhady ACF a PACF se využívají při identifikaci modelu časové řady. V praxi se velmi často vyskytují procesy nestacionární. Nestacionarita procesu může být způsobena střední hodnotou měnící se v čase nebo rozptylem procesu měnícím se v čase. Jestliže po transformaci nestacionárního procesu typu „náhodná procházka“ (tzv. integrovaného procesu) při použití diference  $d$  tého řádu lze výsledný proces popsat stacionárním modelem ARMA ( $p, q$ ), původní integrovaný proces se nazývá autoregresní integrovaný proces klouzavých průměrů řádu  $p, d, q$ , tj. ARIMA ( $p, d, q$ ). (Noskiewičová, ©2008)

Regulační diagram ARIMA (*Autoregressive Integrated Moving Average*) je tedy založen na principu nalezení vhodného modelu časové řady a použití regulačního diagramu na

rezidua modelu (odchylky skutečně naměřené hodnoty od hodnoty vypočtené s použitím modelu). (Noskiewičová, ©2008)

Obecný tvar modelu ARIMA  $(p, d, q)$  je takovýto

$$\Phi_p(B) \cdot \nabla^d \cdot x_t = \Theta_q(B) \varepsilon_t \quad (2.27)$$

kde

$\Phi_p(B) = (1 - \phi_1 B - \phi_2 B^2 \dots - \phi_p B^p)$  je autoregresní polynom  $p$  tého řádu,

$\Theta_q(B) = (1 - \theta_1 B - \theta_2 B^2 \dots - \theta_q B^q)$  je polynom klouzavých průměrů  $q$  tého řádu,

$\nabla$  operátor zpětné difference (tento prvek se zavádí v případě, že modelový proces vykazuje nestacionaritu),

$d$  řád difference,

$t$  čas,

$B$  operátor zpětného posunu  $B \cdot x_t = (x_{t-1})$ ,

$\phi_1, \phi_2, \dots, \phi_p$  parametry autoregresního modelu,

$\theta_1, \theta_2, \dots, \theta_q$  parametry modelu klouzavých průměrů,

$\varepsilon_t$  proměnná zvaná bílý šum, která představuje nepredikovatelnou fluktuaci v datech; má normální rozdělení se střední hodnotou rovnou nule a konstantním rozptylem a její hodnoty jsou nekorelované. (Noskiewičová, ©2008)

Je-li  $\hat{x}_t$  odhad empirické hodnoty  $x_t$  získaný při použití vhodně zvoleného modelu ARIMA, rezidua toho modelu  $e_t = x_t - \hat{x}_t$  se budou chovat jako nezávislé náhodné proměnné pocházející z normálního rozdělení. (Noskiewičová, ©2008)

Nejčastěji se v praxi používají následující ARIMA modely. Uvažujeme model

$$x_t = \xi + \phi x_{t-1} + \varepsilon_t \quad (2.28)$$

kde  $\xi$  a  $\phi$  ( $-1 < \phi < 1$ ) jsou neznámé konstanty a  $\varepsilon_t$  je normálně rozdělená a nezávislá veličina se střední hodnotou rovnou nule a směrodatnou odchylkou  $\sigma$ . Tento model se nazývá autoregresní model 1. řádu a označuje se AR(1). Hodnoty sledovaného znaku

kvality, které jsou navzájem posunuté o  $k$  časových period ( $x_t$  a  $x_{t-k}$ ), mají korelační koeficient  $\phi^k$ . To znamená, že autokorelační funkce ACF by měla exponenciálně klesat. Rozšíříme-li předchozí rovnici do tvaru

$$x_t = \xi + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \varepsilon_t, \quad (2.29)$$

Dostáváme rovnici autoregresního modelu druhého řádu AR(2). Obecně je v autoregresních modelech AR( $p$ ) proměnná  $x_t$  přímo závislá na přechozích hodnotách  $x_{t-1}$ ,  $x_{t-2}$ , atd. Jestliže modelujeme závislost dat pomocí náhodné složky  $\varepsilon_t$ , pak dostáváme modely klouzavých průměrů MA( $q$ ). Model klouzavých průměrů 1. řádu má rovnici:

$$x_t = \mu + \varepsilon_t - \theta \varepsilon_{t-1}. \quad (2.30)$$

V tomto modelu je nenulová korelace pouze mezi dvěma po sobě jdoucími hodnotami  $x_t$  a  $x_{t-1}$  a lze ji vyjádřit následovně:  $\rho_1 = -\theta / (1 + \theta^2)$ . Tomu odpovídá tvar autokorelační funkce ACF. (Arlt et al., ©2002)

Pro modelování praktických úloh se často hodí složený model obsahující jak autoregresní složku, tak složku klouzavých průměrů. Tento model se obecně označuje ARMA( $p, q$ ). Model ARMA 1. řádu, tj. ARMA (1, 1) má rovnici:

$$x_t = \xi + \phi x_{t-1} + \varepsilon_t - \theta \varepsilon_{t-1}. \quad (2.31)$$

Tento model je často vhodný pro chemické a jiné spojité procesy, kde modelem AR(1) lze velmi dobře modelovat mnohé znaky kvality. Náhodnou složkou modelu jsou pak popsány chyby měření, o kterých předpokládáme, že jsou náhodné a nekorelované. V ARMA modelech se předpokládá stacionarita procesu, tzn., že hodnoty sledovaného znaku kvality se pohybují kolem stabilní střední hodnoty. Avšak často se v praxi objevují procesy (např. v chemickém průmyslu, kde sledovaný znak kvality  $x_t$  je výstupní veličinou, který není žádnou regulací udržován na cílové hodnotě), kde hodnoty sledované veličiny „utíkají“. Pak je vhodné modelovat procesy pomocí vhodného modelu s operátorem zpětné difference  $\nabla$ , např. modelem ARIMA (0, 1, 1), jehož rovnice je:

$$x_t = x_{t-1} + \varepsilon_t - \theta \varepsilon_{t-1}. \quad (2.32)$$

Modely ARIMA vyhlížejí odlišně od Shewhartova modelu ( $x_t = \mu + \varepsilon_t$  pro  $t = 1, 2, \dots$ ). Jestliže však do rovnice  $x_t = \xi + \phi x_{t-1} + \varepsilon_t$  dosadíme za  $\phi = 0$  nebo v rovnici

$x_t = \mu + \varepsilon_t - \theta\varepsilon_{t-1}$  za  $\theta = 0$ , dostaneme Shewhartův model procesu. Dalším důležitým krokem při využití ARIMA modelů pro SPC je volba vhodného regulačního diagramu.

Je-li testováním reziduí prokázáno, že nejsou autokorelovaná a pocházejí z normálního rozdělení je možné pomocí nich ověřit, zda proces je či není statisticky stabilní. Protože rozsah výběru  $n=1$  (původní empirické hodnoty  $x_t$  byly zjištěny u každé vyráběné jednotky), nabízí se na prvním místě dvojice regulačních diagramů pro individuální hodnoty a klouzavé rozpětí. Poloha střední hodnoty  $CL$  a hodnoty horní a dolní regulační meze ( $UCL, LCL$ ) se u diagramu ARIMA pro individuální hodnoty stanoví ze vztahu

$$CL = \bar{e} (\cong 0) \quad (2.33)$$

$$UCL = \bar{e} + \frac{3}{1,128} \bar{R}_{kl} \quad (2.34)$$

$$LCL = \bar{e} - \frac{3}{1,128} \bar{R}_{kl} \quad (2.35)$$

kde

$\bar{e}$  je průměrná hodnota reziduí,

$\bar{R}_{kl}$  je průměrné klouzavé rozpětí.

Hodnoty  $CL$ ,  $UCL$  a  $LCL$  v diagramu pro klouzavé rozpětí se určí takto

$$CL = \bar{R}_{kl} \quad (2.36)$$

$$UCL = 3,267 \cdot \bar{R}_{kl} \quad (2.37)$$

$$LCL = 0 \quad (2.38)$$

Pro zvýšení citlivosti regulačních diagramů ARIMA na menší odchylky se doporučuje použít oboustranný regulační diagram CUSUM s rozhodovacím intervalem  $\pm H$  nebo klasický diagram EWMA, oba aplikované na rezidua daného modelu.

Sleduje-li se na jednom produktu  $m$  znaků kvality současně, je možné aplikovat na rezidua  $m$  modelů ARIMA Hotellingův diagram  $T^2$  nebo diagramy CUSUM či EWMA pro vícerozměrné proměnné. (Noskievičová, ©2008), (Malliaris, 1982)

### 2.4.1 Hledání modelu ARIMA

Základní algoritmus hledání modelu ARIMA lze definovat takto:

1. Posouzení stacionarity časové řady z hlediska rozptylu a případná stabilizace časové řady z hlediska rozptylu prostřednictvím vhodné linearizující transformace.
2. Posouzení stacionarity z hlediska střední hodnoty procesu.
3. Stanovení a ověření řádu diferencování (tj. stacionarizace řady z hlediska střední hodnoty).
4. Volba modelu, popř. modelů AR, MA nebo ARMA pro stacionarizovanou časovou řadu, a to na základě hledání podobnosti ve tvaru výběrové ACF a PACF a teoretických ACF a PACF pro známé modely (tento krok zahrnuje stanovení řádu modelů  $p$  a  $q$ ).
5. Odhad parametrů zvoleného modelu, popř. zvolených modelů.
6. Ověření vlastností zvoleného modelu (popř. zvolených modelů) testováním jeho (jejich) reziduí. Model je považován za vyhovující, pokud jeho rezidua vykazují konstantní rozptyl (jsou homoskedastická), nejsou autokorelovaná a mají normální rozdělení.
7. Volba nejlepšího modelu (v případě, že vyhovujících modelů je několik, je nutné použít některé z kritérií pro volbu nejlepšího modelu, např. Akaikeho kritérium AIC, Schwartzovo kritérium SC (nejvhodnější je model pro který dané kritérium nabývá minima). (Noskievičová, ©2008), (Arlt et al., ©2002)

## **II. PRAKTICKÁ ČÁST**

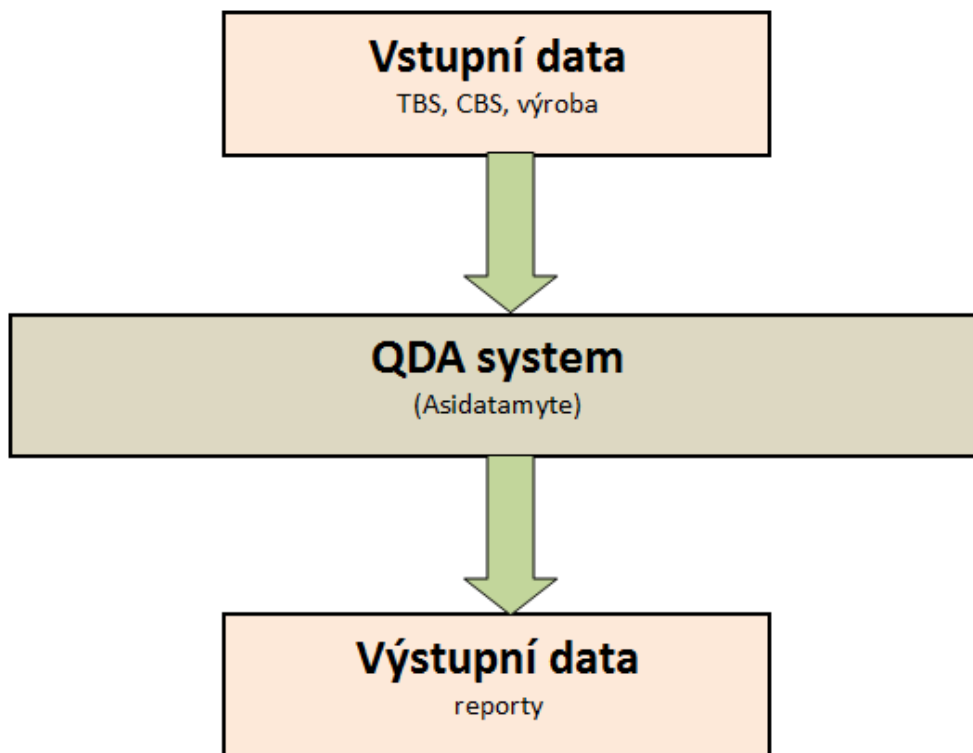
### 3 SYSTÉM AUTOMATICKÉHO SBĚRU DAT PRO MONITOROVÁNÍ KVALITY V PODNIKU CONTINENTAL BARUM, S.R.O. – SPC SOFTWARE

Automatický sběr dat má zajistit kontinuální ukládání a zálohování statistických dat z průběhu řízení výrobních procesů. Tato data jsou nezbytná k řízení, prokazování způsobilosti procesu, analyzování a zlepšování procesu.

HV	
Linka	Měřené veličiny
Lana	Šířka lana, výška lana, průměr lana, hmotnost lana, hmotnost kompletu
VG1	Tloušťka L, tloušťka P, tloušťka <u>Vg_rp</u> L, tloušťka <u>Vg_rp</u> P, šířka PI L, šířka RP L, šířka PI P, šířka RP P, šířka Cel, šířka Roz
VG2	Tloušťka L, tloušťka P, tloušťka <u>Vg_rp</u> L, tloušťka <u>Vg_rp</u> P, šířka PI L, šířka RP L, šířka PI P, šířka RP P, šířka Cel, šířka Roz
VL1	Hmotnost Cel, tloušťka K L, tloušťka R L, tloušťka K P, tloušťka R P, šířka Cel, délka Cel
VL6	Hmotnost Cel, tloušťka K L, tloušťka R L, tloušťka K P, tloušťka R P, šířka Cel, délka Cel
VL8	Hmotnost Cel, tloušťka K L, tloušťka R L, tloušťka K P, tloušťka R P, šířka K, šířka Cel, délka Cel
VL7	Tloušťka B L, tloušťka <u>Pa</u> L, tloušťka B P, tloušťka <u>Pa</u> P, šířka L, šířka P
VL9	Tloušťka B L, tloušťka <u>Pa</u> L, tloušťka B P, tloušťka <u>Pa</u> P, šířka L, šířka P
4V ocel	Hmotnost L, hmotnost S, hmotnost P, tloušťka L, tloušťka S, tloušťka P
4V textil	Hmotnost L, hmotnost S, hmotnost P, tloušťka L, tloušťka S, tloušťka P

Obr. 3.1 Ukázka sběru dat v podniku Continental Barum, s.r.o.

(Zdroj: Vlastní zpracování)



Obr. 3.2 Schématické znázornění vstupu a výstupu systému automatizovaného sběru dat v podniku Continental Barum, s.r.o. (Zdroj: Vlastní zpracování)

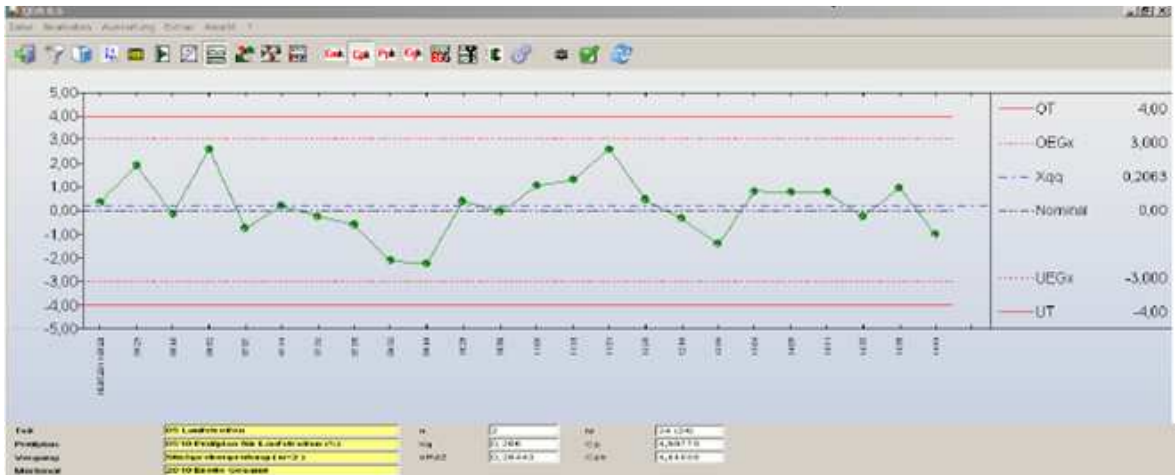
**QDA obecně:**

- Databázový systém pro sběr a vyhodnocování dat.
- Produkt společnosti Asidatamyte – pobočka v ČR.
- V Aachenu systém zaveden od r. 1995.
- QDA – Conti standard SPC nástroj.

**Popis QDA:**

- Procesní charakteristiky jsou měřeny za každý artikl/kód a zadávány do SPC terminálů (počítačů) v absolutních hodnotách.
- Po zadání dat musí operátor kliknout na tlačítko přenést data.
- V případě on-line měření např. profilometry, data budou přenášena automaticky, aby operátor nemohl vybírat data pro vyhodnocování.
- Zadávání dat může být rozděleno na data od operátorů a data od MOK (mezioperační kontrola). Data jsou rozlišována a vyhodnocována odděleně použitím filtru tzn. dvě hodnoty Cpk „výrobní Cpk“ a „MOK Cpk“.
- SPC systém taktéž porovnává zadaná data oproti cíli a limitům a zobrazuje shodu (zelená) nebo neshodu (červeně).
- V případě neshodných měření musí být v QDA systému dokumentována opatření, která byla provedena pro řešení problému.
- V případě neshody (měření mimo tolerance) může být nastaveno automatické odesílání e-mailu pro výrobu nebo procesního inženýra.
- Data lze exportovat do Excelu:
  - SPC systém kalkuluje odchylky od cílových hodnot, procesní data jsou zobrazovány v grafu. Tento graf může být na SPC terminálu kdykoliv zobrazen.
  - Všechny grafy jsou dostupné z každého počítače v podniku s přístupem do QDA.





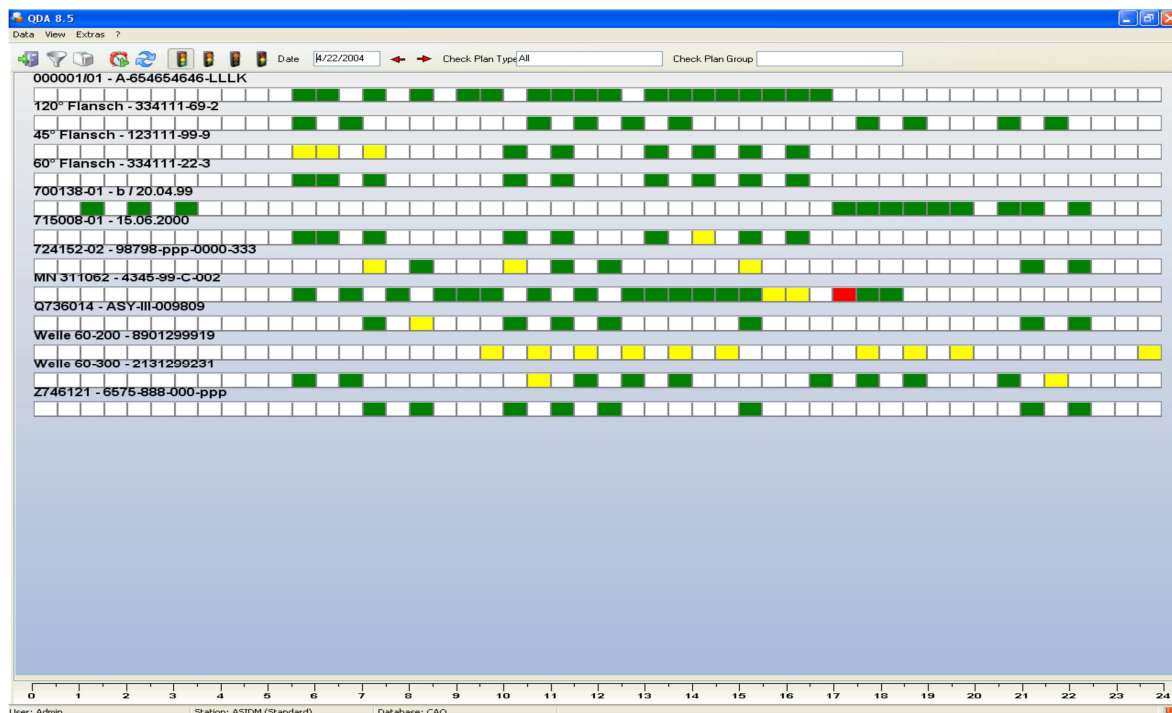
**Obr. 3.3** Možný výstup z QDA v podobě Shewhartova regulačního diagramu

(Zdroj: Výstup ze systému QTREE-SPC)

- Indexy způsobilosti a počty měření jsou rozděleny na výrobu a kvalitu.
- Odchytky, které překračují stanovenou mez, jsou zvýrazněny.
- Standardní časový interval je jeden měsíc, může být změněno.

	Produkt	Merkmal	Produktion		Q- Auditor	
			Cpk <sup>2)</sup>	n	Cpk <sup>3)</sup>	n
06	Gewebelege	0601 Einlagenbreite 193	3,30	516	5,83	60
		0602 Einlagenbreite 194	2,61	177	1,66	21
		0603 Einlagenbreite 196	2,59	420	1,51	48
		0605 Nahtbreite 193	3,52	135	2,05	16
		0606 Nahtbreite 194	3,07	486	3,20	57
		0607 Nahtbreite 196	2,60	161	1,14	19
06	Materialstreifen	0700 Streifenbreite der Wickelbandage	1,88	1325	1,20	160
07	Gürtellage (STC)	0701 1. Breite Gürtellage 185	1,64	537	1,38	45
		0702 2. Breite Gürtellage 185	2,17	528	1,74	45
		0703 1. Breite Gürtellage 186	3,39	600	1,05	45
		0704 2. Breite Gürtellage 186	3,19	591	1,52	45
		0705 1. Breite Gürtellage 187	1,99	576	1,29	60
		0706 2. Breite Gürtellage 187	2,19	573	1,36	54
08	Karkasse	0101 Nahtbreite IS KA mit Mi 6996y	2,28	129	1,79	60
		0102 Nahtbreite IS KA mit Mi 8111y	1,27	30	1,32	6
		0201 Nahtbreite Einlage KA mit Mi 6996y	1,49	129	2,38	60
		0202 Nahtbreite Einlage KA mit Mi 8111y	1,25	30	3,01	6
09	Rohling	0901 1. Breite Gürtellage a. Trommel	2,36	156	1,64	60
		0902 2. Breite Gürtellage a. Trommel	2,17	156	1,52	60
		0903 Breite Wickelbandage a Trommel	2,87	156	1,79	60

**Obr. 3.4** Automatický výpočet indexů způsobilosti procesu (Zdroj: QDA Aachen)



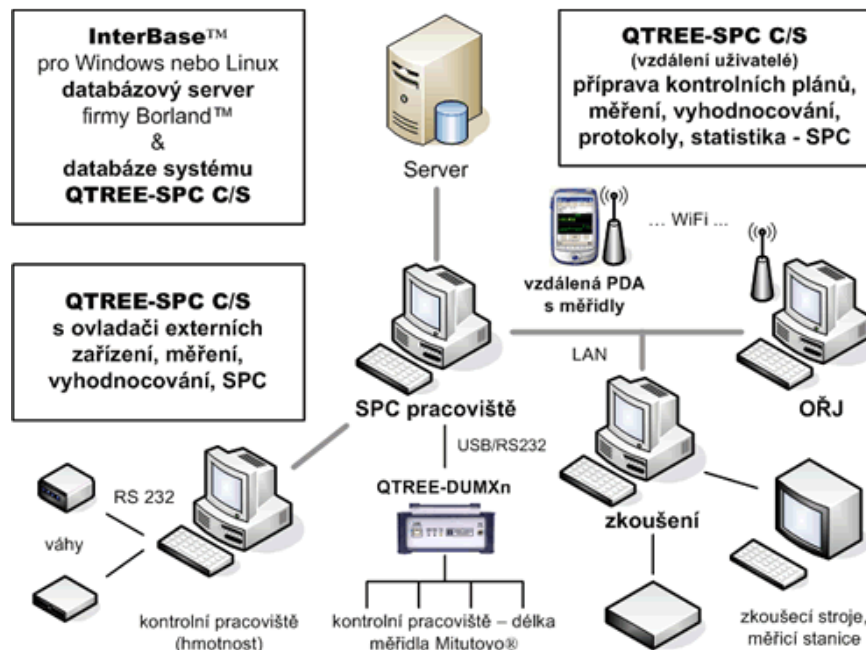
**Obr. 3.5** Ukázka denního reportu, který nabízí jednoduchý přehled kvality po jednotlivých dnech (Zdroj: Výstup ze systému QTREE-SPC)

### 3.1 QTREE-SPC C/S

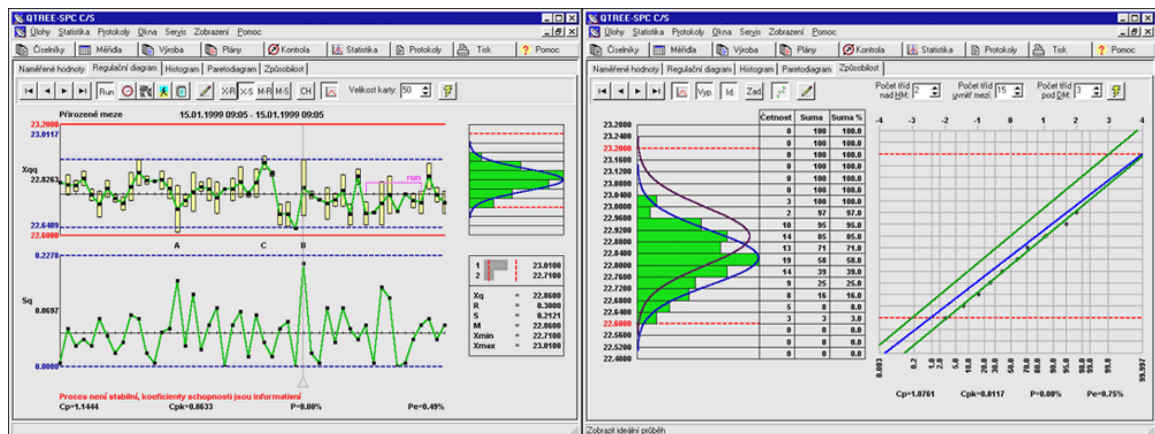
QTREE-SPC C/S je základní součástí produktové řady QTREE CAQ SYSTEMS, která je původním a vlastním softwarovým řešením počítačové podpory řízení kvality - CAQ firmy TŘEŠTÍK. Jedná se o informační systém, který prostřednictvím hierarchické datové struktury kontrolních plánů a kontrolních odběrů zabezpečuje sběr dat<sup>1</sup>, databázové zpracování a vyhodnocování naměřených hodnot rozsáhlou sadou grafických SPC nástrojů pro vizualizaci výsledků statistických analýz. Systém je určen pro pracoviště vstupní, výstupní a mezioperační kontroly, OŘJ a TPV. Je vhodný zejména pro výrobce součástí, dílců a sestav, kteří splňují požadavky norem ISO 9000, ISO 14000, QS9000, VDA 6 a ISO/TS 16949. Softwarové řešení poskytuje jak detailní, tak agregované pohledy na data sledovaných procesů v členění: *Výrobek - Postup - Operace - Výrobní příkaz (dávka,*

<sup>1</sup> QTREE-DUMX Terminal Manager je software pro obsluhu terminálů QTREE-DUMX4-TE a QTREE-DUMX8-TE. Software provádí nastavení terminálů, konfiguruje měřicí úlohy terminálů, přijímá data od terminálů a ukládá je do souborů. Dále exportuje soubory s naměřenými daty do standardních formátů, které lze posléze zpracovat programem MS Excel a statistickým software různých výrobců.

dodávka) - Kontrolní plán - Měřitelné parametry - Atributivní parametry. Z technologického hlediska se jedná se o aplikaci v prostředí Windows (95, 98, XP, NT, 2000, 2003, Vista 32/64bit, 7 32/64-bit) využívající client-server databázovou technologii a on-line komunikaci s externím měřícím zařízením.



Obr. 3.6 Systém QTREE-SPC C/S pro sběr a vyhodnocování procesních dat  
(Zdroj: <http://trestik.cz>)



Obr. 3.7 Možné grafické výstupy ze systému QTREE-SPC C/S  
(Zdroj: <http://trestik.cz/sites/default/files/images/>)

**Základní vlastnosti:**

- Databázové řešení využívá technologii client-server u síťových i lokálních verzí. Standardně je dodáván databázový server InterBase (Windows 95/98/NT/XP/2000, Unix, Solaris, Linux) od firmy Borland. Díky této nejmodernější databázové

technologii, založené na transakčním zpracování, je zajišťována vysoká bezpečnost databáze jak proti neoprávněnému zásahu, tak proti narušení konzistence z důvodu výpadku počítače nebo počítačové sítě.

- Databázová technologie InterBase je firmou Borland uvolněna jako Open Source produkt (zcela zdarma pro libovolnou platformu a libovolný počet klientů).
- Přístupová práva uživatelů jsou přidělována detailně k jednotlivým tabulkám databáze. V důsledku toho je například možné definovat uživatele, který může prohlížet a zpracovávat data, nikoliv však měnit definice kontrolních plánů a naměřené hodnoty.
- Data sledovaných procesů jsou prostřednictvím kontrolních plánů organizována v důmyslně navržené hierarchické struktuře.

### 3.2 Ovladače a externí zařízení

Všechny varianty systému QTREE (QTREE-SPC C/S, QTREE-LAB C/S a QTREE-TEST C/S) jsou vybaveny funkcemi, které zajišťují interaktivní vstup naměřených hodnot z externích měřicích zařízení, připojených prostřednictvím rozhraní RS232 k počítači. Vstupní panely, které jsou určeny pro zadávání naměřených hodnot jsou navrženy tak, že v každém okamžiku, kdy je očekáváno zadání hodnoty z klávesnice, je možný i přímý vstup hodnoty z měřidla.

Komunikaci mezi měřidly a programy zajišťují samostatné moduly - programové můstky - ovladače externích zařízení. Ke komunikaci mezi ovladači a volajícími programy je využívána technologie DDE (Dynamic Data Exchange). Každý ovladač je realizován jako samostatný "DDE Server" a volající program, který využívá služeb ovladače je realizován jako "DDE client". Komunikace mezi ovladačem a klientským programem probíhá obousměrně. To znamená, že klientský program není jenom v roli příjemce naměřené hodnoty, ale poskytuje ovladači informace nutné pro komunikaci s externím zařízením (například číslo vstupního kanálu u vícekanálových zařízení). Z tohoto důvodu nemají ovladače zcela univerzální využití. Není například možné prostřednictvím těchto ovladačů přijímat hodnoty do tabulky programu MS Excel. Pro vstup dat z měřicích zařízení do tabulky programu MS Excel jsou k dispozici jiné, speciálně k tomu účelu vytvořené programy. V každém okamžiku mohou být v systémech QTREE aktivní maximálně tři ovladače, přičemž každý ovladač může v případě vícekanálových zařízení obsluhovat prakticky libovolný počet kanálů.

## 4 ŘEŠENÍ PROBLEMATIKY MONITOROVÁNÍ AUTOKORELOVANÝCH PROCESŮ V SOUČASNÉ PRAXI

V následujícím textu se zaměřím na problematiku monitorování autokorelovaných procesů v současné praxi, kde na případových studiích budu demonstrovat aplikaci regulačních diagramů pro časové řady při řešení reálných procesů, se kterými jsem se setkal ve své praxi, a které jsem osobně řešil.

### 4.1 Případová studie sestrojení dynamického diagramu EWMA

Proces válcování hliníku za studena má být statisticky regulován. Analýzou procesu bylo prokázáno, že jakost povrchu plechu je významně ovlivňována viskozitou chladicí kapaliny. Bylo proto rozhodnuto měřit jednou denně viskozitu chladicí kapaliny. Proces byl nastaven na počáteční hodnotu  $6,00 \text{ m}^2 \cdot \text{s}^{-1}$ . Výsledky 50 měření (v  $\text{m}^2 \cdot \text{s}^{-1}$ ) jsou uvedeny v tabulce 4.1., sl.  $x_k$ .

#### Krok 1:

Z analýzy regulačních diagramů ( $x_j, R$ ) (viz obr. 4.2), které byly zprvu vybrány pro statistickou regulaci uvedeného procesu, plyne, že proces nemá konstantní ani střední hodnotu a velké množství bodů mimo regulační meze vede k podezření, že data jsou závislá. Pomocí autokorelační funkce (viz obr. 4.3) bylo dále prokázáno, že mezi naměřenými hodnotami viskozity existuje pozitivní autokorelace.

Pro tuto situaci je vhodným nástrojem statistické regulace dynamický diagram EWMA.

#### Krok 2:

Optimální hodnota parametru  $\lambda$  byla stanovena následovně:

- a) dle vztahu (2.5) byly postupně vypočítány hodnoty  $\text{EWMA}_{k-1}$  (za  $k$  bylo do vzorce dosazeno  $k - 1$ ) pro  $\lambda = 0,1, 0,2, 0,3, 0,4, 0,5, 0,51, 0,55, 0,56, 0,57, 0,58, 0,59, 0,6, 0,7, 0,8, 0,9$ . Výpočet  $\text{EWMA}_{k-1}$  se inicializuje tak, že první hodnota  $\text{EWMA}$  ( $\text{EWMA}_0$ ), tj. odhad 1. hodnoty viskozity  $\hat{y}_1$ , se rovná požadované hodnotě viskozity  $Y_0 = 6$ . Hodnoty  $\text{EWMA}_{k-1}$  pro  $k = 1, 2, \dots, 51$  při  $\lambda = 0,56$  jsou uvedeny v následující tabulce, sl.  $\text{EWMA}_{k-1}$ . Tyto hodnoty byly určeny následovně:

Např.

$$\text{EWMA}_1 = (1 - 0,56) \cdot 6,00 + 0,56 \cdot 6,128 = 6,07168,$$

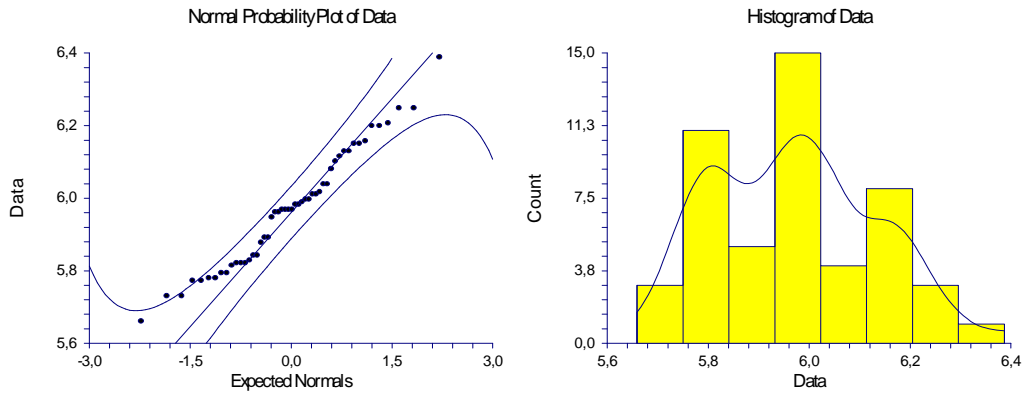
$$EWMA_2 = (1 - 0,56) \cdot 6,07168 + 0,56 \cdot 6,198 = 6,142419.$$

- b) Pro každý soubor hodnot  $EWMA_{k-1}$  (pro různé hodnoty  $\lambda$  – viz bod a) byly dle vztahu (2.5) stanoveny chyby predikce  $e_k$  pro  $k = 1, 2, \dots, 50$ . Hodnoty chyby predikce  $e_k$  pro  $\lambda = 0,56$  jsou uvedeny v tabulce 4.1, sl.  $e_k$ .
- c) Pro každý soubor hodnot  $EWMA_{k-1}$  (pro různé hodnoty  $\lambda$  – viz bod a) byla dále vypočtena hodnota sumy čtverců chyb predikce  $\sum_{k=1}^{50} e_k^2$  (viz tab. 4.2).

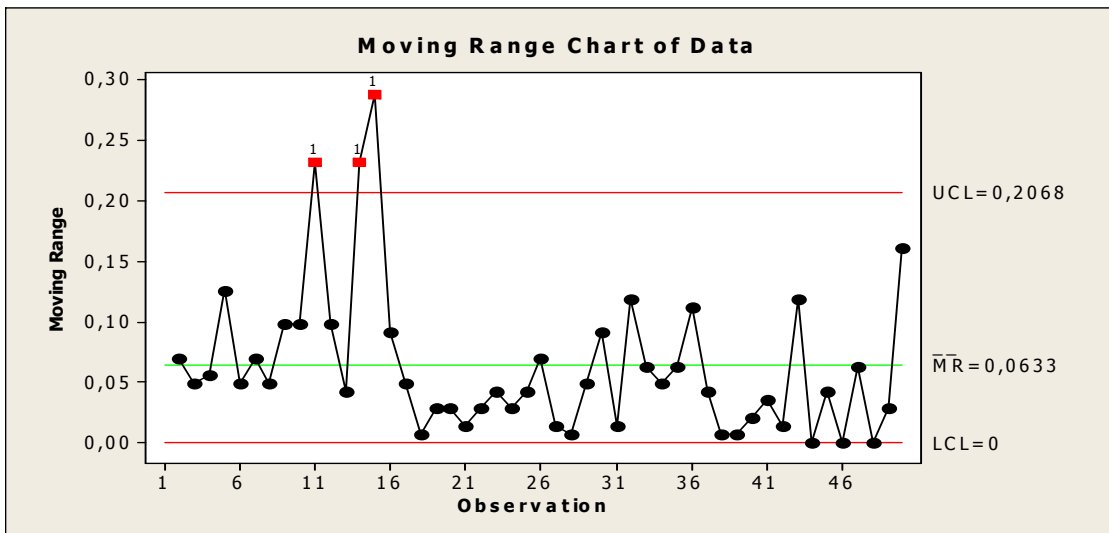
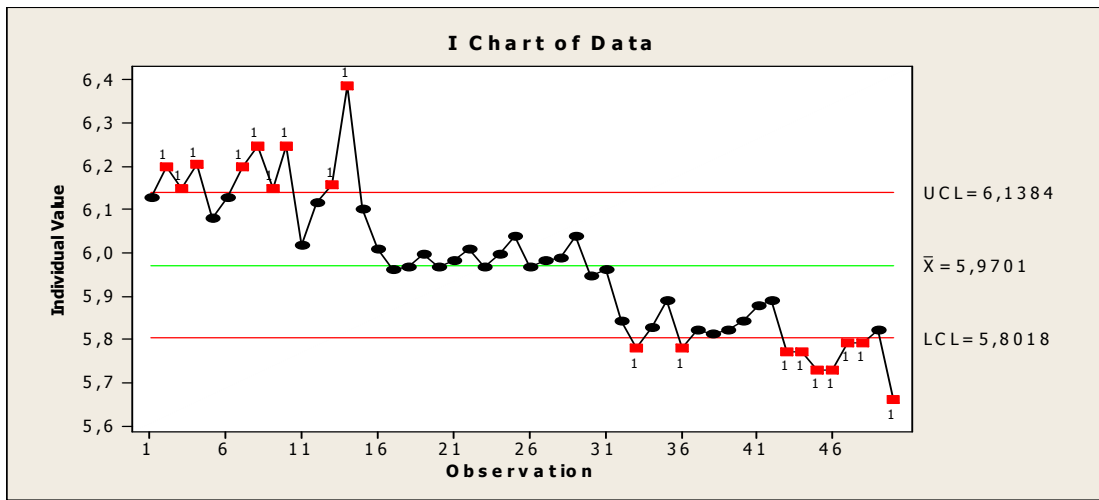
**Tab. 4.1** Vstupní údaje vypočtené hodnoty  $EWMA_{k-1}$  a  $e_k$  pro  $\lambda = 0,56$

(Zdroj: Vlastní zpracování)

$k$	$x_k$	$EWMA_{k-1}$	$e_k$	$k$	$x_k$	$EWMA_{k-1}$	$e_k$
1	6,128	6	0,128	26	5,967	6,0155767	-0,0486
2	6,198	6,07168	0,12632	27	5,981	5,9883737	-0,0074
3	6,149	6,1424192	0,00658	28	5,988	5,9842444	0,00376
4	6,205	6,1461044	0,0589	29	6,037	5,9863476	0,05065
5	6,079	6,179086	-0,1001	30	5,946	6,0147129	-0,0687
6	6,128	6,1230378	0,00496	31	5,96	5,9762337	-0,0162
7	6,198	6,1258166	0,07218	32	5,841	5,9671428	-0,1261
8	6,247	6,1662393	0,08076	33	5,778	5,8965028	-0,1185
9	6,149	6,2114653	-0,0625	34	5,827	5,8301413	-0,0031
10	6,247	6,1764847	0,07052	35	5,890	5,8283822	0,06162
11	6,016	6,2159733	-0,2000	36	5,778	5,8628881	-0,0849
12	6,114	6,1039882	0,01001	37	5,820	5,8153508	0,00465
13	6,156	6,1095948	0,04641	38	5,813	5,8179543	-0,0005
14	6,387	6,1355817	0,25142	39	5,820	5,8151799	0,00482
15	6,100	6,276376	-0,1764	40	5,841	5,8178792	0,02312
16	6,009	6,1776054	-0,1686	41	5,876	5,8308268	0,04517
17	5,960	6,0831864	-0,1232	42	5,890	5,8561238	0,03388
18	5,967	6,014202	-0,0472	43	5,771	5,8750945	-0,1041
19	5,995	5,9877689	0,00723	44	5,771	5,8168016	-0,0458
20	5,967	5,9918183	-0,0248	45	5,729	5,7911527	-0,0622
21	5,981	5,9779201	0,00308	46	5,729	5,7563472	-0,0273
22	6,009	5,9796448	0,02936	47	5,792	5,7410328	0,05097
23	5,967	5,9960837	-0,0291	48	5,792	5,7695744	0,02243
24	5,995	5,9797968	0,0152	49	5,820	5,7821327	0,03787
25	6,037	5,9883106	0,04869	50	5,659	5,8033384	-0,1443
				51		5,7225089	



**Obr. 4.1** Exploratorní analýza ukazuje, že data pocházejí z normálního rozdělení  
(Zdroj: Vlastní zpracování)



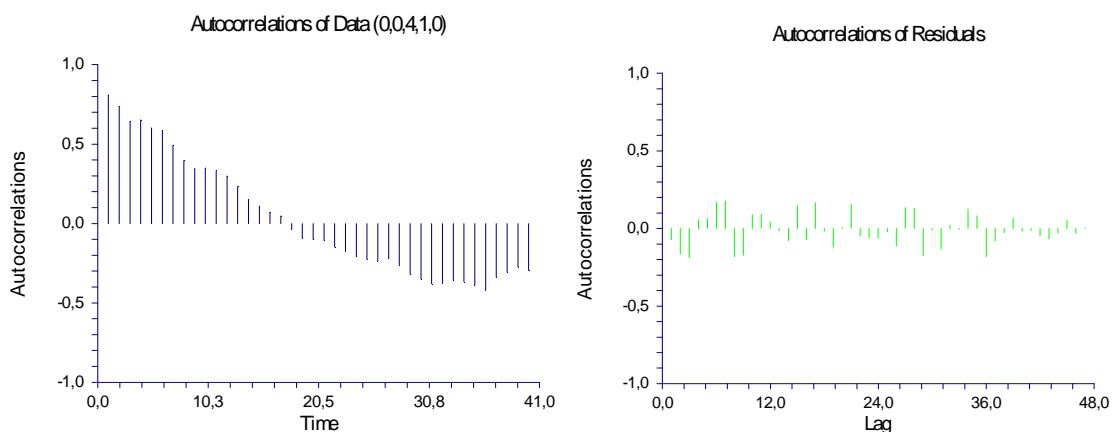
**Obr. 4.2** Regulační diagramy  $x_j$ ,  $R_{kl}$  pro naměřené hodnoty viskozity  
(Zdroj: Vlastní zpracování)

**Tab. 4.2** Hodnoty sumy čtverců chyb predikce  $\sum_{k=1}^{50} e_k^2$  pro různé hodnoty  $\lambda$

(Zdroj: Vlastní zpracování)

$\lambda$	$\sum e_k^2$	$\lambda$	$\sum e_k^2$
0,1	0,65124111	0,57	0,35154
0,2	0,44405202	0,58	0,35159
0,3	0,3833032	0,59	0,35169
0,4	0,36066646	0,6	0,35183
0,5	0,3526581	0,7	0,35551
0,51	0,35231971	0,8	0,3631
0,55	0,35157957	0,9	0,37514
<b>0,56</b>	<b>0,35153294</b>		

Z tabulky je patrné, že minimální hodnoty této sumy bylo dosaženo při  $\lambda = 0,56$ , tzn., že  $\lambda_{\text{opt}} = 0,56$ .



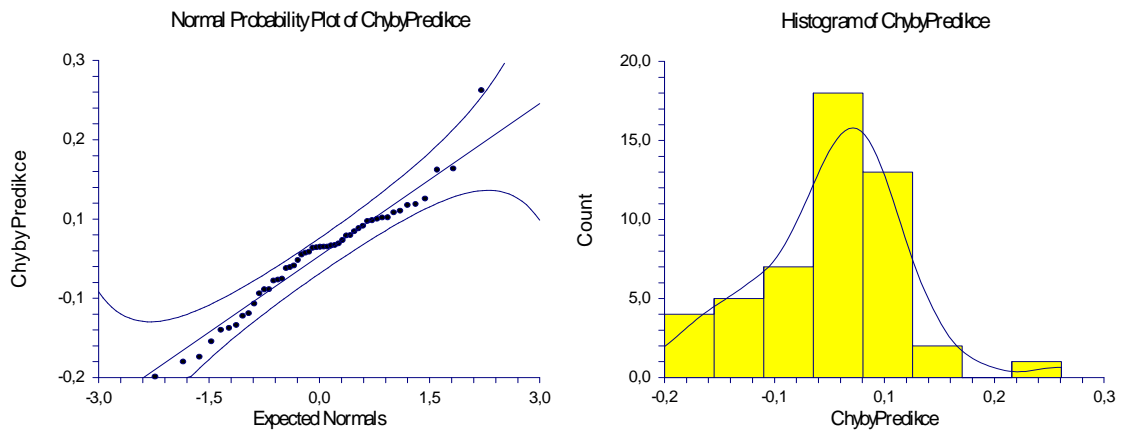
**Obr. 4.3** Autokorelační funkce (vlevo), autokorelace reziduí (vpravo)

(Zdroj: Vlastní zpracování)

### Krok 3:

Pomocí Shapiro-Wilkova a Anderson-Darlingova testu byla ověřena a potvrzena normalita chyb predikce  $e_k$  pro  $\lambda_{\text{opt}}$ .





**Obr. 4.4** Exploratorní analýza ukazuje na normalitu chyb predikce  $e_k$  pro  $\lambda_{opt}$   
(Zdroj: Vlastní zpracování)

Náhodnost těchto reziduí byla ověřena a potvrzena Durbin-Watsonovým testem náhodnosti.

#### Durbin-Watson Test For Serial Correlation

Parameter	Value	Did the Test Reject
		H0: Rho(1) = 0?
Durbin-Watson Value	1,8262	
Prob. Level: Positive Serial Correlation	0,2207	No
Prob. Level: Negative Serial Correlation	0,6808	No

Výsledky potvrzují správnost volby dynamického diagramu EWMA pro statistickou regulaci daného procesu a správnost stanovení optimální hodnoty parametru  $\lambda$ .

#### Krok 4:

Odhad  $\sigma_p$  vypočteme dosazením do vztahu (2.10). Pak

$$\hat{\sigma}_p = \sqrt{\frac{0,35153294}{49}} = 0,0847003.$$

#### Krok 5:

Hodnoty bodů střední čáry  $CL_k$  a regulačních mezí dynamického EWMA diagramu pro  $\lambda_{opt}$  byly stanoveny pomocí vztahů (2.7) – (2.10). Např.:

$$CL_1 = 6, UCL_1 = 6 + 3 \cdot 0,0847003 = 6,2541, LCL_1 = 6 - 3 \cdot 0,0847003 = 5,745899086.$$

Hodnoty  $CL_k$ ,  $UCL_k$ ,  $LCL_k$  pro všechny hodnoty  $k$  jsou uvedeny v tabulce 4.3.

**Krok 6:**

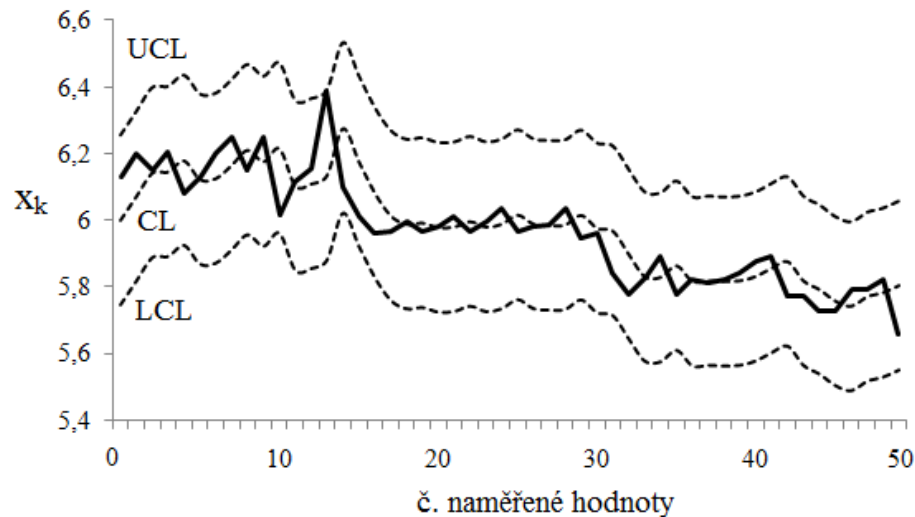
Z uvedených hodnot  $CL_k$ ,  $UCL_k$  a  $LCL_k$  byl sestrojen diagram a do něj byly vyneseny naměřené hodnoty viskozity  $x_k$  (viz obr. 4.5).

**Tab. 4.3** Hodnoty střední čáry a regulačních mezí dynamického diagramu EWMA pro  $\lambda_{opt.} = 0,56$  (Zdroj: Vlastní zpracování)

$k$	$CL_k$	$UCL_k$	$LCL_k$	$k$	$CL_k$	$UCL_k$	$LCL_k$
1	6	6,2541	5,7459	26	6,01558	6,26968	5,76148
2	6,07168	6,32578	5,81758	27	5,98837	6,24247	5,73427
3	6,14242	6,39652	5,88832	28	5,98424	6,23835	5,73014
4	6,1461	6,40021	5,892	29	5,98635	6,24045	5,73225
5	6,17909	6,43319	5,92499	30	6,01471	6,26881	5,76061
6	6,12304	6,37714	5,86894	31	5,97623	6,23033	5,72213
7	6,12582	6,37992	5,87172	32	5,96714	6,22124	5,71304
8	6,16624	6,42034	5,91214	33	5,8965	6,1506	5,6424
9	6,21147	6,46557	5,95736	34	5,83014	6,08424	5,57604
10	6,17648	6,43059	5,92238	35	5,82838	6,08248	5,57428
11	6,21597	6,47007	5,96187	36	5,86289	6,11699	5,60879
12	6,10399	6,35809	5,84989	37	5,81535	6,06945	5,56125
13	6,10959	6,3637	5,85549	38	5,81795	6,07206	5,56385
14	6,13558	6,38968	5,88148	39	5,81518	6,06928	5,56108
15	6,27638	6,53048	6,02228	40	5,81788	6,07198	5,56378
16	6,17761	6,43171	5,9235	41	5,83083	6,08493	5,57673
17	6,08319	6,33729	5,82909	42	5,85612	6,11022	5,60202
18	6,0142	6,2683	5,7601	43	5,87509	6,1292	5,62099
19	5,98777	6,24187	5,73367	44	5,8168	6,0709	5,5627
20	5,99182	6,24592	5,73772	45	5,79115	6,04525	5,53705
21	5,97792	6,23202	5,72382	46	5,75635	6,01045	5,50225
22	5,97964	6,23375	5,72554	47	5,74103	5,99513	5,48693
23	5,99608	6,25018	5,74198	48	5,76957	6,02368	5,51547
24	5,9798	6,2339	5,7257	49	5,78213	6,03623	5,52803
25	5,98831	6,24241	5,73421	50	5,80334	6,05744	5,54924

**Krok 7:**

Dynamický diagram EWMA pro uvedený proces signalizuje významnou odchylku u 14. hodnoty. Je třeba najít její příčinu a přijmout nápravné opatření.



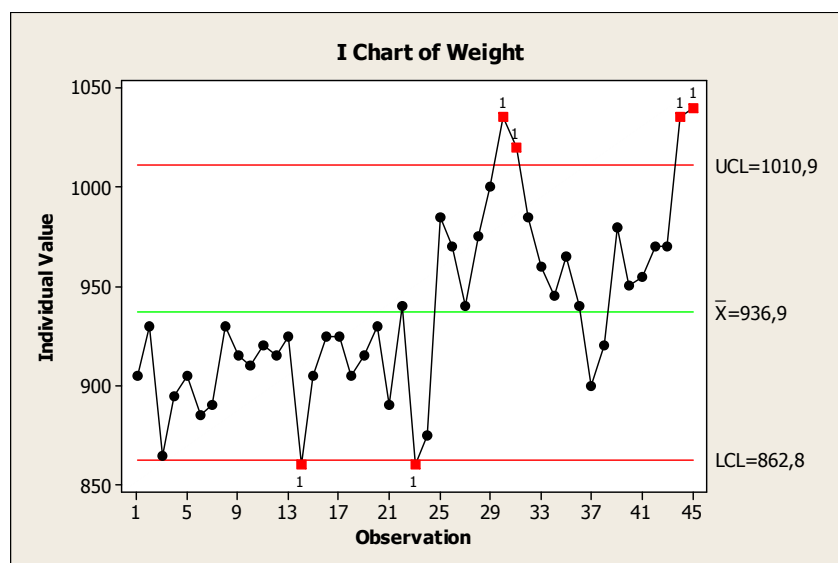
**Obř. 4.5** Dynamický diagram EWMA pro naměřené hodnoty viskozity

(Zdroj: Vlastní zpracování)

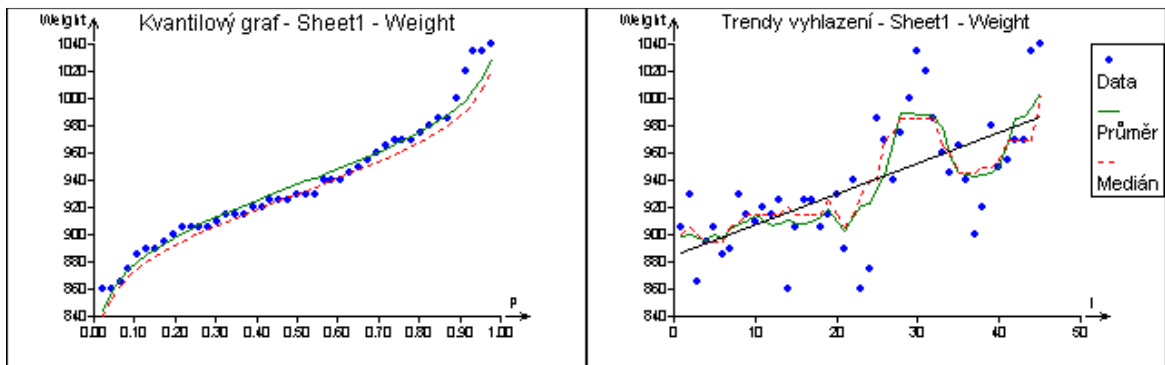
## 4.2 Případová studie aplikace použití regulačního diagramu EWMA

V této případové studii máme naměřená data z chemické analýzy škodlivých zplodin, která musí být monitorována pomocí regulačního diagramu s cílem detekovat odchylky od statistické stability.

**Řešení:** Protože se jedná o individuální hodnoty bez opakování, bylo by formálně možné použít diagramu “x-individual”. Jako centrální linii lze použít průměr dat,  $CL = 936,9$ , a regulační meze lze vypočítat z klouzavého rozpětí,  $UCL = 1010,9$ ;  $LCL = 862,8$ . Po konstrukci mezí a doplnění dat získáme následující diagram.

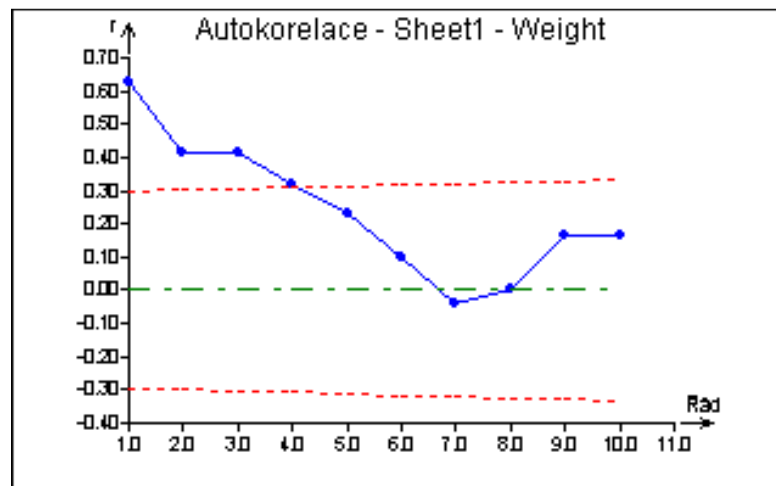


**Obř. 4.6** Diagram “x individual” (Zdroj: Vlastní zpracování)



**Obr. 4.7** Kvantilový graf potvrzuje normalitu dat (vlevo). Data s vyznačeným trendem a vyhlazením (vpravo) (Zdroj: Vlastní zpracování)

Několikeré překročení regulačních mezí, porušuje podmínky dané zvláštními případy uvnitř regulačních mezí, které je nutno hledat v porušení některého z předpokladů o datech pro konstrukci regulačních diagramů. V předchozím grafu je zřejmý statisticky významný lineární trend znázorněný přímkou a kolísání hodnot, které je typické pro autokorelovaná data. Tento předpoklad potvrzuje i graf autokorelace na následujícím obrázku, kde autokorelační koeficient prvního, druhého, třetího a čtvrtého řádu je statisticky významný. Lze tedy konstatovat, že statistickým modelem našich dat s největší pravděpodobností není nezávislé normální rozdělení s konstantní střední hodnotou, jak to vyžaduje konstrukce Shewhartova diagramu "x individual".



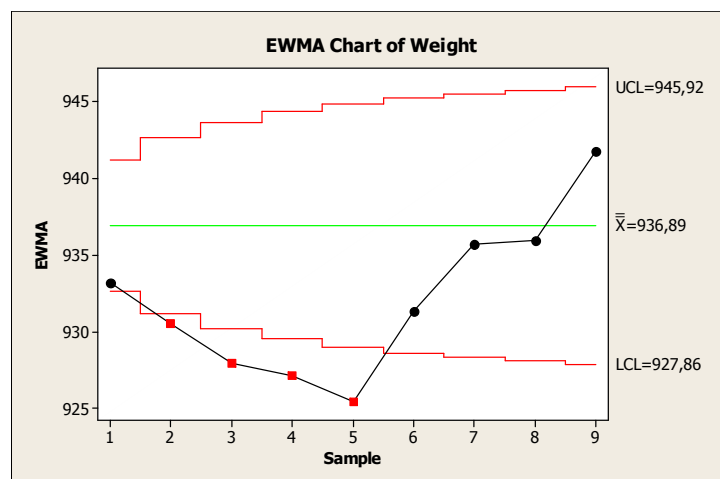
**Obr. 4.8** Graf autokorelace s mezemi významnosti korelačních koeficientů (Zdroj: Vlastní zpracování)

## Autocorrelations of Weight (0,0,4,1,0)

Lag	Correlation	Lag	Correlation	Lag	Correlation	Lag	Correlation
1	0,627773	11	0,109513	21	-0,272486	31	-0,133205
2	0,411944	12	0,106562	22	-0,230590	32	-0,089071
3	0,410377	13	0,246299	23	-0,174390	33	-0,076887
4	0,315451	14	0,279330	24	-0,254361	34	-0,062777
5	0,233693	15	0,133365	25	-0,265942	35	-0,092027
6	0,100511	16	0,023592	26	-0,252097	36	-0,098975
7	-0,043601	17	0,030366	27	-0,245672	37	-0,103388
8	0,000636	18	-0,026339	28	-0,207322	38	-0,163957
9	0,167105	19	-0,035892	29	-0,199735	39	-0,138539
10	0,165032	20	-0,074043	30	-0,192246	40	-0,114650

Significant if  $|Correlation| > 0,298142$

Regulačním diagramem, který připouští autokorelaci i nekonstantnost střední hodnoty, může být například některý model časové řady jako autoregrese *AR*, klouzavý průměr *MA*. V praxi často je používán dynamický diagram EWMA s jednokrokovou predikcí na následujícím obrázku, případně rezidua tohoto diagramu.



**Obr. 4.9** Dynamický diagram EWMA s jednokrokovou predikcí

(Zdroj: Vlastní zpracování)

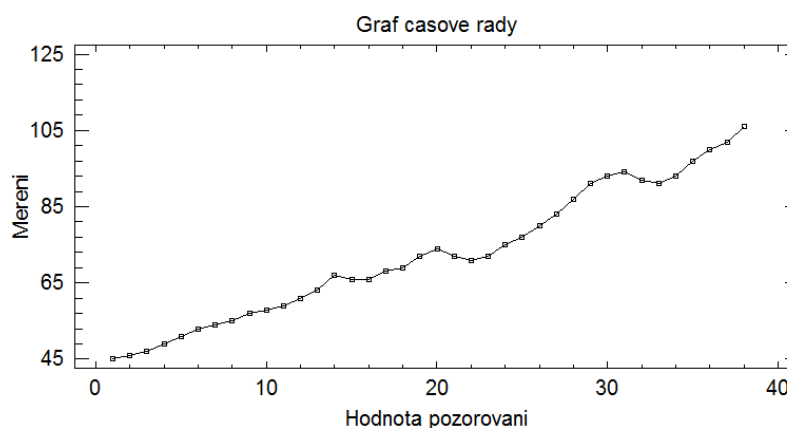
Tento poslední diagram ukazuje čtyři hlavní problémy: bod 2, 3, 4 a bod 5, na něž je třeba se zaměřit při interpretaci. Ostatní 2 problémy detekované diagramem "x-individual" jsou zřejmě důsledkem závislosti a trendu v datech. Tento závěr však neznamená, že závislostí nebo trendu v datech se není třeba zabývat. Pokud by se podařilo autokorelaci technologicky vysvětlit a odstranit, dosáhli bychom snížení variability, zvýšení indexu způsobilosti, a tím zlepšení jakosti o faktor  $(1 - r^2)$ , tedy v tomto případě zhruba o 25 %, protože autokorelační koeficient 1. řádu je  $r_1 = 0,627$ .

### 4.3 Praktická případová studie – Použití regulačních diagramů ARIMA, EWMA a CUSUM na reálném procesu

Pro měření uvedené v tabulce níže pro měření šířky patního pásku u bočnice najdeme vhodný model a na jeho základě určíme regulační diagram citlivý na změnu střední hodnoty. Následující výstupy jsou zpracovány v programu Statgraphics Centurion.

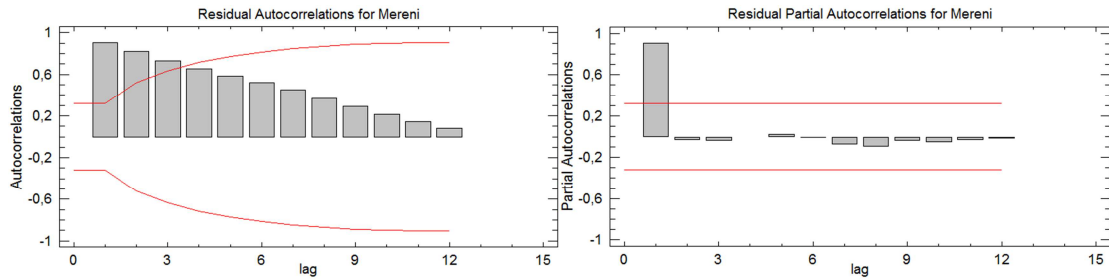
**Tab. 4.4** Šířka patního pásku u bočnice v  $1 \cdot 10^{-1}$  mm. (Zdroj: Vlastní zpracování)

$t$	$x_t$	$t$	$x_t$	$t$	$x_t$	$t$	$x_t$
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		

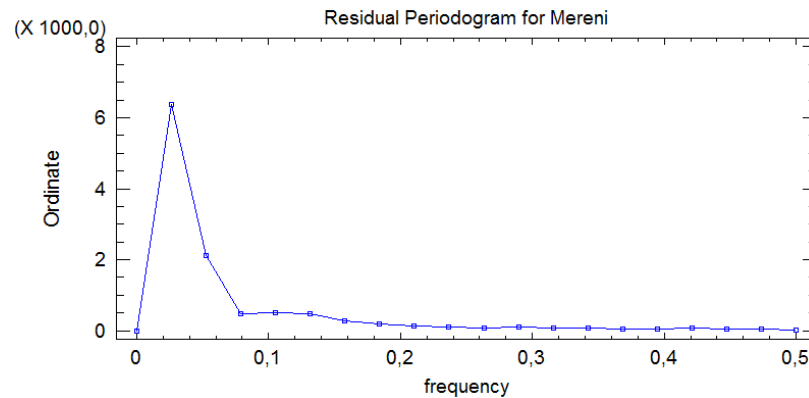


**Obr. 4.10** Znárodnění časové řady měření šířky patního pásku u bočnice  
(Zdroj: Vlastní zpracování)

Graf časové řady je znázorněn na předchozím obrázku 4.10. z obrázku je patrné, že časová řada vykazuje nestacionaritu, což je také potvrzeno tvarem ACF, PACF a periodogramem.

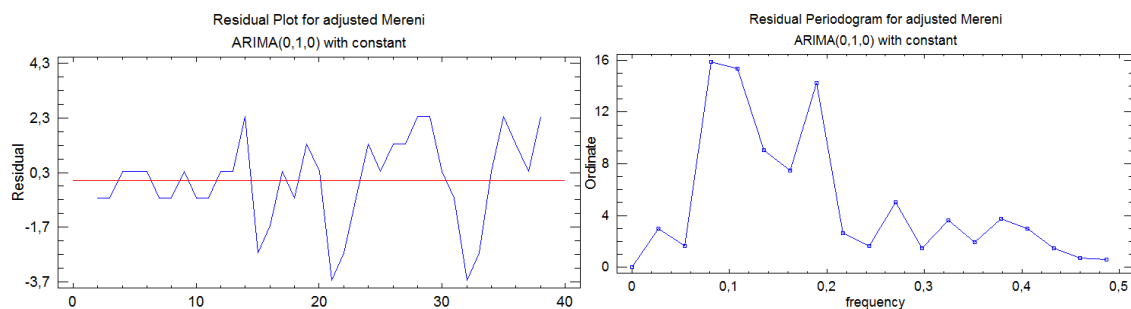


**Obr. 4.11** ACF a PACF původní časové řady (Zdroj: Vlastní zpracování)



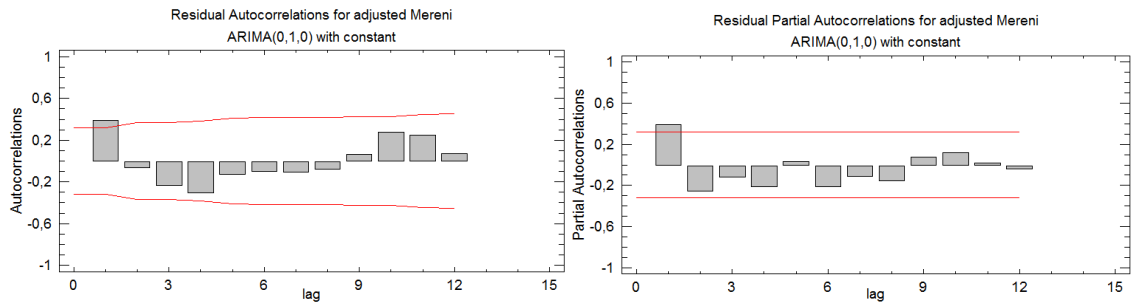
**Obr. 4.12** Periodogram původní časové řady (Zdroj: Vlastní zpracování)

Jelikož je patrné, že hodnoty ACF klesají pomalu, první hodnota u ACF i PACF se blíží hodnotě 1 a periodogram má významný vrchol v nulové frekvenci, lze předpokládat, že časová řada je typu  $I(1)$ . Na základě této skutečnosti ji tedy budeme stacionarizovat prvními diferencemi. Tento závěr je potvrzen také grafem reziduí a reziduálním periodogramem (obr. 4.13) a stejně tak hodnoty reziduálního ACF a PACF grafu (obr. 4.14) jejichž hodnoty leží, až na první, uvnitř tolerančních mezí.



**Obr. 4.13** Graf reziduí a reziduální periodogram modelu ARIMA(0,1,0)

(Zdroj: Vlastní zpracování)



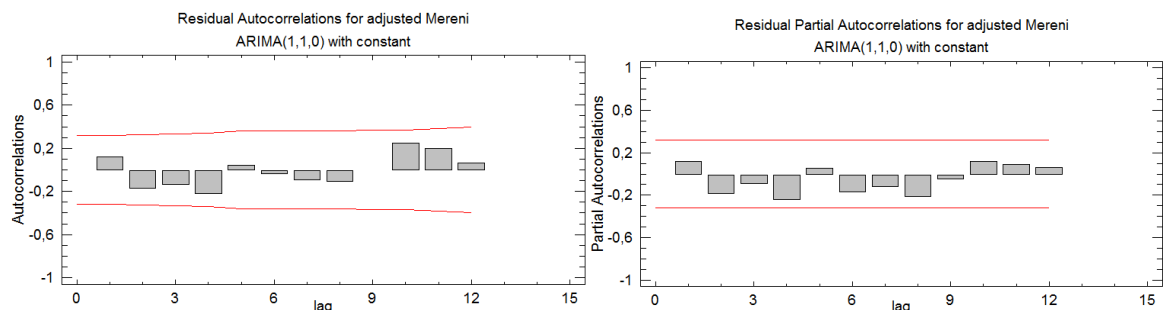
**Obr. 4.14** Reziiduální ACF a PACF modelu  $ARIMA(0,1,0)_c$

(Zdroj: Vlastní zpracování)

**Tab. 4.5** Výstupní tabulka pro model  $ARIMA(0,1,0)_c$  (Zdroj: Vlastní zpracování)

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>	<b>ARIMA Model Summary</b>		
<i>Statistic</i>	<i>Period</i>	<i>Period</i>	<i>Parameter</i>	<i>Estimate</i>	<i>Std. Error</i>
RMSE	1,60236		Mean	1,64865	0,263427
MAE	1,22863		Constant	1,64865	
MAPE	1,62859				
ME	3,07262E-15		Backforecasting: yes		
MPE	-0,106827		Estimated white noise variance = 2,56757 with 36 degrees of freedom		
			Estimated white noise standard deviation = 1,60236		
<b>Tests for Randomness of residuals</b>					
Box-Pierce Test					
Test based on first 12 autocorrelations					
Large sample test statistic = 18,8926					
P-value = 0,0911537					

Reziiduální grafy ACF i PACF mají první hodnoty statisticky významně odlišné od nuly. Nelze proto jednoznačně určit, zda je model třeba rozšířit o část  $AR(1)$  či  $MA(1)$ . Odhadneme tedy nejprve model  $ARIMA(1,1,0)_c$ . Výsledky jsou uvedeny v tab. 4.6.



**Obr. 4.15** Reziiduální ACF a PACF modelu  $ARIMA(1,1,0)_c$

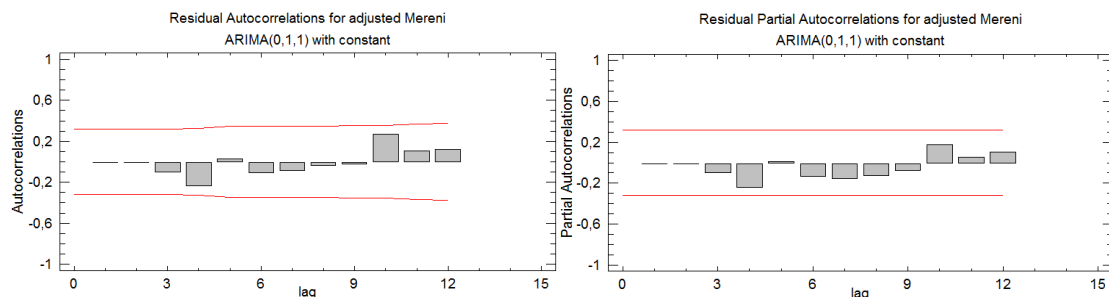
(Zdroj: Vlastní zpracování)



I když reziduální grafy ACF a PACF naznačují, že nesystematická složka je typu bílého šumu a rovněž  $p$  hodnota Boxova-Pierceova testu je relativně vysoká (0,597248), provedeme i odhad modelu  $ARIMA(0,1,1)_c$ , jeho výsledky obsahuje tab. 4.7.  $P$  hodnota v tomto případě je tedy  $\geq \alpha$  (0,05), tudíž nezamítáme  $H_0$  o neautokorelovanosti reziduí.

**Tab. 4.6** Výstupní tabulka modelu  $ARIMA(1,1,0)_c$  (Zdroj: Vlastní zpracování)

Nonseasonal differencing of order: 1			<b>ARIMA Model Summary</b>	
Forecast model selected: $ARIMA(1,1,0)$ with constant			<i>Parameter</i>	<i>Estimate</i>
Number of forecasts generated: 4			<i>Std. Error</i>	<i>t</i>
Number of periods withheld for validation: 0			<i>P-value</i>	
<i>Statistic</i>	<i>Estimation Period</i>	<i>Validation Period</i>	AR(1)	0,407307
RMSE	1,48723		Mean	1,66965
MAE	1,09575		Constant	0,989589
MAPE	1,46969		Backforecasting: yes	
ME	0,0162096		Estimated white noise variance = 2,21333 with 35 degrees of freedom	
MPE	-0,0476283		Estimated white noise standard deviation = 1,48773	
<b>Tests for Randomness of residuals</b>			Number of iterations: 1	
Box-Pierce Test				
Test based on first 12 autocorrelations				
Large sample test statistic = 9,26715				
P-value = 0,597248				



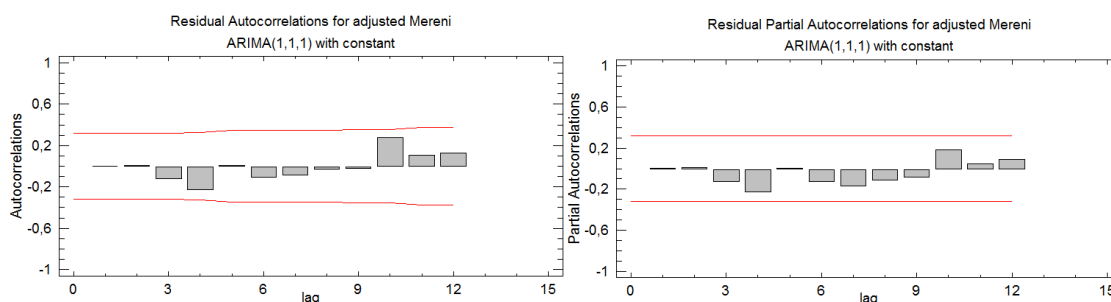
**Obr. 4.16** Reziduální ACF a PACF modelu  $ARIMA(0,1,1)_c$

(Zdroj: Vlastní zpracování)

Reziduální ACF a PACF i u tohoto modelu naznačují, že nesystematická složka je typu bílého šumu,  $p$  hodnota Boxova-Pierceova testu (0,794106) je dokonce vyšší než u modelu s částí AR(1). Pro úplnost odhadneme ještě model obsahující obě složky, tj. AR(1) a MA(1). Tyto modely pak porovnáme z pohledu interpolačních a extrapolacních kritérií a po stránce vyšší  $p$  hodnoty Boxova-Pierceova testu pro neautokorelovanost nesystematické složky.

Tab. 4.7 Výstupní tabulka modelu  $ARIMA(0,1,1)_c$  (Zdroj: Vlastní zpracování)

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				
Statistic	Estimation Period	Validation Period		
RMSE	1,43305			
MAE	0,9854			
MAPE	1,32419			
ME	0,00350342			
MPE	-0,0689866			
Parameter	Estimate	Std. Error	t	P-value
MA(1)	-0,569874	0,147994	-3,85066	0,000480
Mean	1,67905	0,362276	4,63471	0,000048
Constant	1,67905			
Backforecasting: yes				
Estimated white noise variance = 2,054 with 35 degrees of freedom				
Estimated white noise standard deviation = 1,43318				
Number of iterations: 3				
Tests for Randomness of residuals				
Box-Pierce Test				
Test based on first 12 autocorrelations				
Large sample test statistic = 7,06119				
P-value = 0,794106				

Obr. 4.17 Reziiduální ACF a PACF modelu  $ARIMA(1,1,1)_c$ 

(Zdroj: Vlastní zpracování)

I když reziiduální ACF a PACF, které jsou na obr. 4.17 ukazují, že rezidua mají nesystematický charakter a  $p$  hodnota Boxova-Pierceova testu je relativně vysoká (0,701229), z  $p$  hodnot  $t$  testů parametrů modelu je zřejmé, že část  $AR(1)$  do modelu nepatří. Tento  $t$  test testuje hypotézu  $H_0$  o statistické nevýznamnosti parametru modelu, oproti  $H_A$  o jeho statistické významnosti.

Nejen tento závěr však svědčí ve prospěch modelu  $ARIMA(0,1,1)_c$ . Ve srovnání s modelem  $ARIMA(1,1,0)_c$  tento model vede k vyšší  $p$  hodnotě Boxova-Pierceova testu, a k nižším hodnotám interpolačních a extrapoláčních kritérií, tedy MSE, MAE, MAPE, ME a MPE.

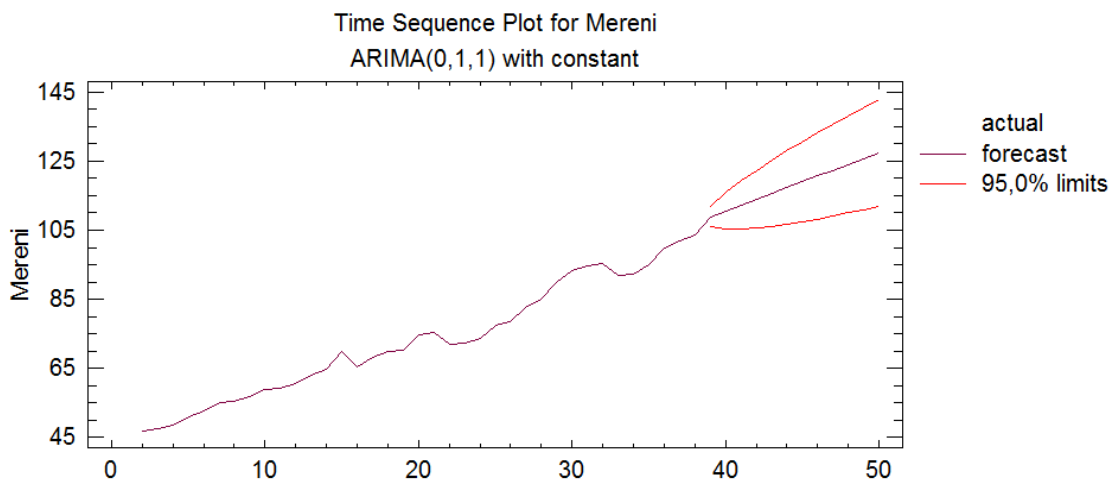
Tab. 4.8 Výstupní tabulka pro model ARIMA(1,1,1)<sub>c</sub> (Zdroj: Vlastní zpracování)

Nonseasonal differencing of order: 1			ARIMA Model Summary				
Forecast model selected: ARIMA(1,1,1) with constant			Parameter	Estimate	Std. Error	t	P-value
Number of forecasts generated: 4			AR(1)	-0,0807671	0,312105	-0,258782	0,797364
Number of periods withheld for validation: 0			MA(1)	-0,635226	0,246258	-2,57952	0,014390
Statistic	Estimation Period	Validation Period	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				

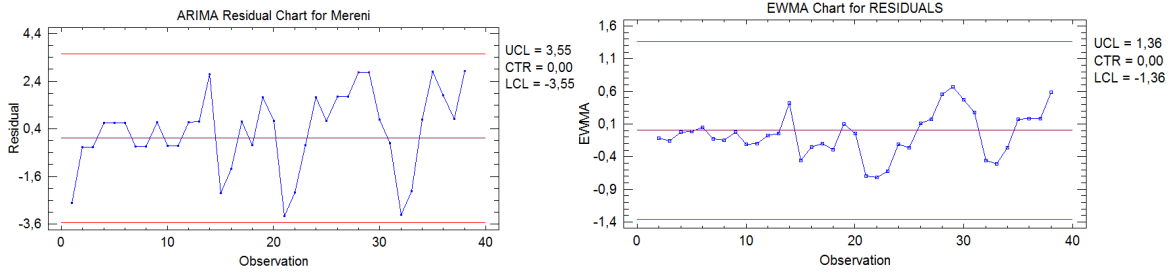
Tests for Randomness of residuals		
Box-Pierce Test		
Test based on first 12 autocorrelations		
Large sample test statistic = 7,25441		
P-value = 0,701229		

Po porovnání  $p$  hodnot parametrů  $c = \mu$  (0,000048) a  $\theta_1$  (0,000480) s hladinou významnosti  $\alpha$  (0,05) u provedeného  $t$  testu z tab. 4.7 je zřejmé, že oba parametry jsou nenulové a model s odhadnutými parametry má tvar  $(1-B)y_t = 1,67905 + (1 - (0,569874)B)a_t$ , lze jej vyjádřit také jako  $y_t = 1,67905 + y_{t-1} + a_t + 0,569874a_{t-1}$ .



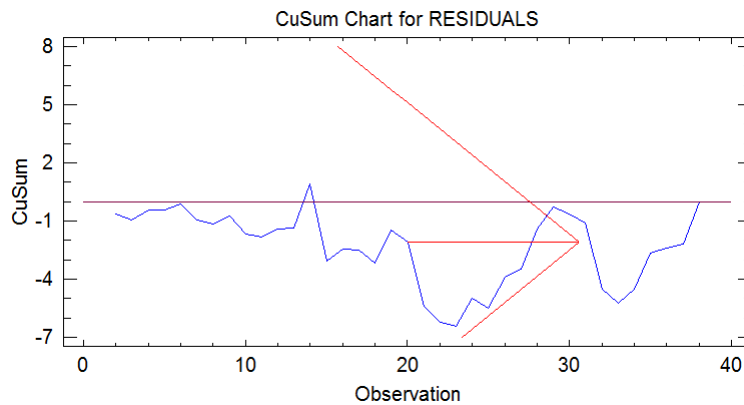
Obr. 4.18 Časová řada s předpověďmi (Zdroj: Vlastní zpracování)

Z následujících regulačních diagramů je zřejmé, že proces můžeme považovat za statisticky stabilní (žádný bod se nenachází mimo regulační meze), tudíž tyto meze mohou být použity pro ověření statistické stability tohoto procesu v následujícím období.



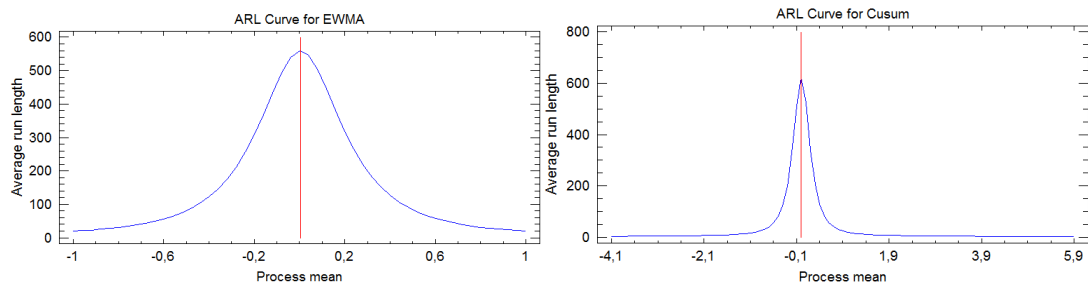
**Obr. 4.19** Regulační diagram pro individuální hodnoty aplikovaný na rezidua zvoleného ARIMA modelu (vlevo) a regulační diagram EWMA pro rezidua identifikovaného modelu ARIMA s parametrem  $\lambda = 0,2$  (vpravo).

(Zdroj: Vlastní zpracování)



**Obr. 4.20** Regulační diagram CUSUM (Zdroj: Vlastní zpracování)

Z regulačního diagramu CUSUM je patrné, že V-masko identifikovala počátek neustále se zvětšující heteroskedasticity kolem 23. měření. Předchozí regulační diagramy ukazují shodný bod změny ve volatilitě této časové řady a díky zachycenému průběhu neustále se zvětšující heteroskedasticity naznačují rostoucí trend této časové řady.



**Obr. 4.21** Křivky ARL pro změny střední hodnoty u diagramů EWMA (vlevo) a CUSUM (vpravo) (Zdroj: Vlastní zpracování)

Z předchozích grafů ARL je patrné, že citlivost pro malé posuny procesní úrovně lépe detekuje diagram CUSUM než diagram EWMA.

## 5 NÁVRH REGULAČNÍHO APARÁTU PRO AUTOMATIZOVANÝ SBĚR DAT VE VÝROBNÍCH PROCESECH

V této kapitole budu srovnávat jednak aparát, navržený Atienzou a Tangem (2002) a aparát navržený Montgomerym a Friedmanem (1989) s klasickým Shewhartovým regulačním diagramem po stránce výkonnosti v simultánním monitorování střední hodnoty, rozptylu a autokorelační struktury sériově korelovaných procesů. Dále se zaměřím na studium chování regulačního diagramu  $\lambda_{LS,t}$  navrženým Atienzou et al. (1998) používaného pro detekci úrovně posunu v procesu AR(1). Výkonnost tohoto aparátu bude vyhodnocena a srovnána s Alvanovým a Robertsovým aparátem SCC (diagram zvláštních příčin – modifikovaný Shewhartův regulační diagram pro rezidua). Pro srovnání výkonnosti těchto regulačních diagramů použiji simulaci Monte Carlo implementovanou v programovém prostředí jazyka R, která se nachází v knihovně `spc`.

### 5.1 Simultánní monitorování střední hodnoty, rozptylu a autokorelační struktury sériově korelovaných procesů

Statistické metody v řízení procesů pro monitorování sériově korelovaných nebo autokorelovaných procesů získaly velkou pozornost v literatuře, týkající se statistického řízení kvality. Zaměření většiny výzkumných studií je však na detekci posunu střední hodnoty procesu. Detekce změny ve variabilitě a autokorelační struktuře časové řady v důsledku některých zvláštních příčin variability ovlivňující systém je často přehlížena. V ideálním případě je potřeba udržovat tři regulační diagramy, aby se účinně detekovaly změny v procesu: jeden pro detekci změny střední hodnoty, jeden pro změnu variability a jeden pro změnu autokorelační struktury (MVAS – Mean, Variance and Autocorrelation Structure) časové řady. Takový přístup může být k realizaci docela náročný. Jednou alternativou je vytvořit monitorovací systém, který má dobrou citlivost v simultánní detekci změn v MVAS časové řady. Příkladem takového přístupu je výběrový autokorelační diagram (SACC – Sample Autocorrelation Chart), který navrhl Montgomery a Friedman (1989). Bohužel bylo zjištěno, že SACC má nízkou citlivost při detekci posunu střední hodnoty a rozptylu. Jako alternativa bude v této kapitole ukázán návrh nového monitorovacího aparátu, který je založený na vlastnostech Gaussovské stacionární časové řady, viz. (Atienza et al., 2002). Ve srovnání s SACC, navrhovaný aparát je jednodušší k implementaci a je více citlivý na změny v MVAS v časové řady.

Techniky statistického řízení procesů (SPC) se zaměřují na co nejrychlejší zjištění přítomnosti vnějších nebo vymežitelných příčin variability, které ovlivňují daný proces. Tyto zvláštní příčiny variability obvykle mají za následek nežádoucí odchylky důležitých charakteristik výrobku od požadované nebo cílové hodnoty. SPC postupy spoléhají především na předpoklad, že pokud je proces ovlivněn pouze inherentními nebo vymežitelnými příčinami variability, bude proces produkovat měření, která jsou v čase nezávislá a se stejným rozdělením. Jakákoliv odchylka od těchto i.i.d. vlastností (independent and identically distributed random variables – nezávislé náhodné veličiny pocházející ze stejného rozdělení), která jsou tím pádem pod kontrolou (v regulačních mezích), je obvykle interpretována jako signalizace přítomnosti zvláštních příčin variability. Tento předpoklad byl základem klasických SPC technik.

Díky nedávným pokrokům v technologii výroby, jsou tyto techniky náročné na použitelnost, co se týká i.i.d. konceptu, tedy procesu, který je pod kontrolou (Montgomery a Friedman, 1989). Například, v nepřetržitém chemickém zpracování a automatizovaných výrobních prostředích, jsou procesní měření často sériově korelovaná, i když tam nejsou žádné zvláštní příčiny variability ovlivňující systém. Tyto autokorelační porušení předpokládají nezávislost procesních dat ve stavu pod kontrolou (tedy v regulačních mezích) v mnoha procedurách při sestavování regulačních diagramů. Takové porušení má významný vliv na výkon klasických SPC technik (Harris a Ross, 1991), (Alwan, 1992).

Nejpopulárnější přístup k monitorování procesů se sériově korelovanými pozorováními je modelovat inherentní autokorelaci procesních měření s použitím modelu autoregresivních klouzavých průměrů (ARMA). Pro detekci změn v procesu, jsou použita rezidua zvoleného modelu. Je-li tento model vhodně vybrán a dobře odhadnut, tato rezidua mají přibližné i.i.d. vlastnosti. Díky tomuto předpokladu pak už můžeme uplatnit tradiční SPC diagramy pro monitorování reziduí. Jakákoliv odchylka reziduí od i.i.d. vlastností naznačuje změnu v procesu, která nastala v důsledku vymežitelné příčiny ve variabilitě.

Rozvoj SPC technik pro monitorování autokorelace procesů dosáhl značné pozornosti v literatuře týkající se inženýrství kvality. Zaměření je hlavně na detekci posunu střední hodnoty procesu. Detekce změny ve variabilitě (volatilitě) a autokorelační struktury časové řady, které jsou také důležitými ukazateli přítomnosti procesních změn, musí pak v důsledku vymežitelné příčiny ovlivňovat systém, je však často přehlížena.

Dooley a Kapoor (1990) navrhli monitorování změn ve střední hodnotě, variabilitě a autokorelační struktuře (MVAS) procesních měření pomocí simultánního znázorňování CUSUM,  $\chi^2$  a autokorelačního diagramu na rezidua odhadnutého modelu z daných procesních dat. Yourstone a Montgomery (1989) poukázali na to, že autokorelační funkce (ACF) „bude detekovat změny v autokorelační struktuře stejně dobře, jako posuny střední hodnoty a variability dat procesu kvality, v reálném čase“. Navrhli monitorování v reálném čase prvních  $m$  reziduálních autokorelací vypočítaných pro  $n$  nejnovějších procesních měření. Tento graf nazývají výběrovým autokorelačním diagramem nebo (SACC – Sample Autocorrelation Chart). Atienza a kol. (1997) zkoumali průměrnou délku série (ARL – Average Run Length – střední doba do výskytu nevyhovující hodnoty) u SACC. Poukázali na to, že i když SACC může detekovat změny ve střední hodnotě a variabilitě řady, zdaleka není tak dobrý, jako Shewhartův regulační diagram (SCC) použitý na rezidua. Jedinou výhodou SACC je ta, že může zjistit změny v autokorelační struktuře řady lépe než SCC. Takže, nelze spoléhat jen na SACC pro simultánní monitorování změny v MVAS řady. V této kapitole bude ukázán přístup navržený Atienzou a Tangem, který bude sloužit jako alternativa k SACC. Ve srovnání s SACC, navrhovaná procedura, která je založena na vlastnostech stacionárního Gaussova procesu, je mnohem jednodušší na implementaci a je citlivá na posuny v MVAS časové řady. Ačkoliv navrhovaný postup může být implementován pro obecné ARMA případy, bude analýza zaměřena na důležitý případ procesu AR (1).

### 5.1.1 Návrh nového přístupu pro detekci posunu v MVAS

Budeme se řídit obvyklým předpokladem, že když je proces stabilní, posloupnost procesních měření  $\{x_t\}$  může být modelována stacionárním Gaussovským ARMA( $p$ ,  $q$ ) procesem.

$$\phi_p(B)x_t = \phi_0 + \theta_q(B)\varepsilon_t, \quad (5.1)$$

kde  $\{x_t\}$  je stacionární časová řada reprezentující procesní měření,

$$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \quad (5.2)$$

je autoregresivní polynom řádu  $p$ ,

$$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \quad (5.3)$$

je klouzavý průměr polynomu řádu  $q$ ,  $B$  je operátor zpětného posunu a  $\{\varepsilon_t\}$  je posloupnost Gaussovských nezávislých náhodných chyb, s nulovou střední hodnotou a konstantním rozptylem  $\sigma_\varepsilon^2$ . Bez ztráty na obecnosti budeme předpokládat, že počáteční hodnota  $\varphi_0$  procesních dat procesu pod kontrolou je nulová.

Podle definice, pokud  $\{x_t, t = 1, 2, 3, \dots\}$  je stacionární Gaussův proces, potom  $\{x_t\}$  je striktně stacionární, neboť pro všechna  $n \in \{1, 2, 3, \dots\}$  a pro všechna  $h, t_1, t_2, \dots \in \mathbb{Z}$ , náhodné vektory  $(x_{t_1}, \dots, x_{t_n})'$  a  $(x_{t_1+h}, \dots, x_{t_n+h})'$  mají stejnou střední hodnotu a stejnou kovarianční matici a z toho důvodu i stejné rozdělení (Brockwell a Davis, 1987, s. 13).

Nechť  $\{x_t\}$  je stacionární Gaussův proces ARMA. Potom  $\mathbf{x} = (x_{t_1}, \dots, x_{t_n})'$  je vícerozměrné normální rozdělení se střední hodnotou  $\mu_0$  a kovarianční maticí

$$\Sigma_0 = \begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \cdots & \gamma_{n-1} \\ \gamma_1 & \gamma_0 & \gamma_1 & \cdots & \gamma_{n-2} \\ \gamma_2 & \gamma_1 & \gamma_0 & \cdots & \gamma_{n-3} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \gamma_{n-1} & \gamma_{n-2} & \gamma_{n-3} & \cdots & \gamma_0 \end{bmatrix} = \sigma_x^2 = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{n-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{n-2} \\ \rho_2 & \rho_1 & 1 & \cdots & \rho_{n-3} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \rho_{n-1} & \rho_{n-2} & \rho_{n-3} & \cdots & 1 \end{bmatrix}, \quad (5.4)$$

kde  $\gamma_k$  a  $\rho_k$  reprezentují autokovarianci a autokorelaci v řádu zpoždění  $k$ , dle pořadí. Pro proces ARMA(1,1) autokovarianční funkce je dána následovně

$$\gamma_0 = \frac{1 + \theta_1^2 - 2\phi_1\theta_1}{1 - \phi_1^2} \sigma_\varepsilon^2, \quad (5.5)$$

$$\gamma_1 = \frac{(1 - \phi_1\theta_1)(\phi_1 - \theta_1)}{1 - \phi_1^2} \sigma_\varepsilon^2, \quad (5.6)$$

$$\gamma_k = \phi_1 \gamma_{k-1}, \quad k \geq 2, \quad (5.7)$$

přičemž autokorelační funkce je dána

$$\rho_k = \begin{cases} 1, & k = 1, \\ \phi_1 \rho_{k-1}, & k \geq 2. \end{cases} \quad (5.8)$$

Autokovarianční a autokorelační funkce obecného ARMA ( $p, q$ ) procesu lze nalézt v (Box et al., 1994, s. 76). Postupy pro odhad autokovarianční funkce jsou podrobně



projednávány v (Brockwell a Davis, 1987) a (Box et al., 1994, s. 76). Z výše uvedené charakteristiky Gaussovského procesu ARMA, víme, že

$$H_n = \mathbf{x}' \Sigma_0^{-1} \mathbf{x} \quad (5.9)$$

má  $\chi^2$  rozdělení s  $n$  stupni volnosti. Tato statistika je citlivá na změny jak ve vektoru středních hodnot, tak i autokovarianční matici. Pokud rozdělíme  $\{x_1, x_2, x_3, \dots\}$  do disjunktní řady všech pozorování délky  $n$  ( $\{\mathbf{x}'_1 = (x_1, x_2, x_3, \dots, x_n), \mathbf{x}'_2 = (x_{n+1}, x_{n+1}, \dots, x_{2n}), \dots\}$ ), můžeme zvolit  $n$  takové dostatečně velké, že korelace mezi  $\{\mathbf{x}'_1, \mathbf{x}'_2, \dots\}$  je zanedbatelná. Proto můžeme aplikovat typický  $\chi^2$  regulační diagram pro monitorování ARMA procesu. Vzhledem k tomu, že tento regulační diagram je analogický s vícerozměrným  $T^2$  regulačním diagramem pro monitorování výrobků nebo procesů s několika korelovanými inherentními znaky kvality, (Montgomery, 2009), (Atienza et al. 1997) může tedy rozšířit působnost na analogii vícerozměrného CUSUM nebo exponenciální vážený klouzavý průměr (EWMA) pro monitorování autokorelovaných procesů.

Při implementaci výše uvedeného  $\chi^2$  regulačního diagramu, po výpočtu statistiky  $H$ , je potřeba čekat na všechny následující  $n$  pozorování před výpočtem další statistiky  $H$ . Právě to může zpomalit detekci procesních změn, zejména je-li  $n$  velké. Jednou z alternativ je výpočet statistiky  $H$  pro nejnovějších  $n$  pozorování, tj. s použitím následující posloupnosti vektoru měření:

$$\mathbf{x}_1^* = (x_1, x_2, \dots, x_n)', \mathbf{x}_2^* = (x_2, x_3, \dots, x_{n+1}'), \mathbf{x}_3^* = (x_3, x_4, \dots, x_{n+2}'), \dots$$

V následujícím textu budeme nazývat diagram  $H$  na základě výše uvedené posloupnosti měření váženým  $H$  diagramem nebo také  $MH$  diagram. Pro tento případ, výsledné statistiky  $H$  již nadále nejsou nezávislé. Proto nemůžeme stanovit regulační meze pomocí stejného přístupu, jako při konstrukci regulačního diagramu  $T^2$ . K určení regulačních mezí můžeme použít simulace, čímž dosáhneme požadované ARL pod kontrolou.

### 5.1.1.1 Vlastnosti a provedení ARL

Monitorovací strategie popsána v předchozí části je samozřejmě Shewhartova koncepce. Pokud  $n = 1$ , lze snadno ověřit, že navrhovaná regulační strategie se stává zvláštním případem Shewhartova regulačního diagramu s modifikovanými regulačními mezemi.

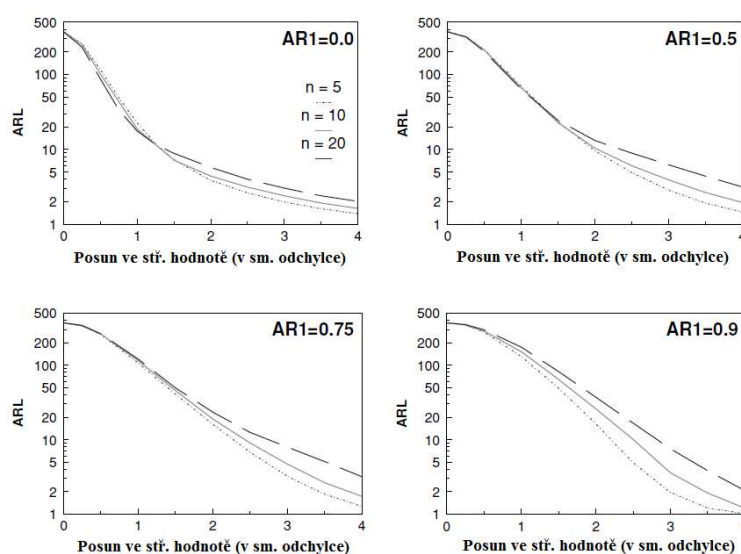
Proto v následujících odstavcích pouze porovnáme ARL vlastnosti navrženého aparátu s podobným Shewhartovým monitorovacím aparátem (SCC a jeho konkurent SACC).

### 5.1.1.2 Citlivost na detekci posunu střední hodnoty

Citlivost diagramu *MH* je závislá na rozsahu výběru  $n$ . Pomocí simulace Monte Carlo, budeme analyzovat vliv  $n$  na schopnost *MH* grafu odhalit změnu ve střední hodnotě u procesu AR(1). Posun ve střední hodnotě je vyjádřen podmínkami  $\sigma_x$ ,

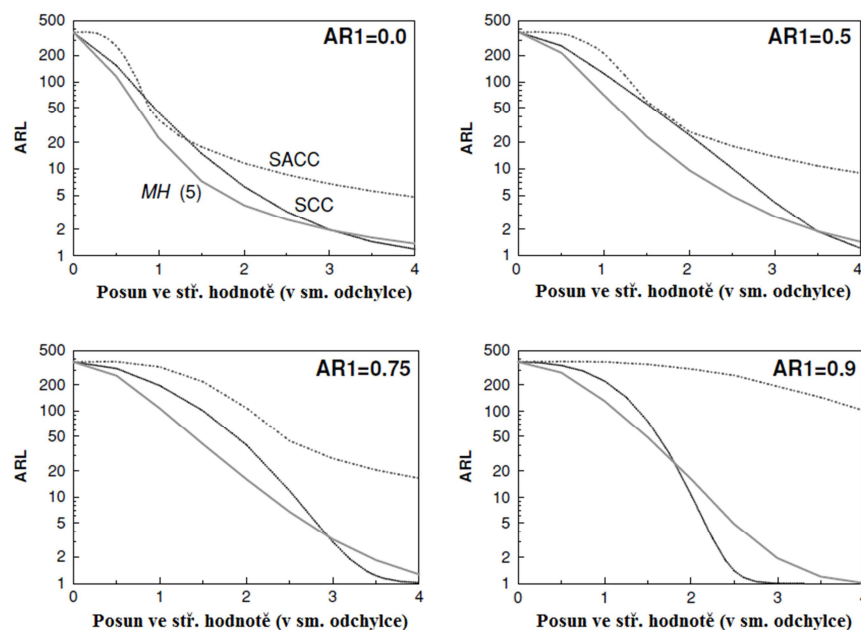
$$\sigma_x = \sqrt{\frac{\sigma_\varepsilon^2}{1 - \phi_1^2}}, \quad (5.10)$$

kde  $\phi_1$  reprezentuje autoregresní parametr. Každý běh simulace je tvořen z 5000 iterací. Požadovaný proces ARMA je generován pomocí inovovaného algoritmu (Brockwell a Davis, 1987), zatímco posloupnost nezávislých normálně rozdělených náhodných veličin je generována pomocí statistické knihovny `spc` jazyka R. Pro účely srovnání jsou následně všechny regulační meze v *MH* diagramu zvoleny tak, aby ARL pod kontrolou byla přibližně 370. Vliv  $n$  na citlivost *MH* diagramu při detekci posunu střední hodnoty je znázorněno na následujícím obrázku. Je zřejmé, že pokud je autokorelace nízká, pro detekci malých změn ve střední hodnotě potřebujeme velké  $n$ . Nicméně, když je proces vysoce pozitivně autokorelován, už při  $n$  blízko 5 dává dobrou citlivost na detekci pro malé i velké posuny ve střední hodnotě.



**Obr. 5.1** Vliv změny  $n$  na citlivost *MH* diagramu při detekci posunu střední hodnoty (Zdroj: Vlastní zpracování)

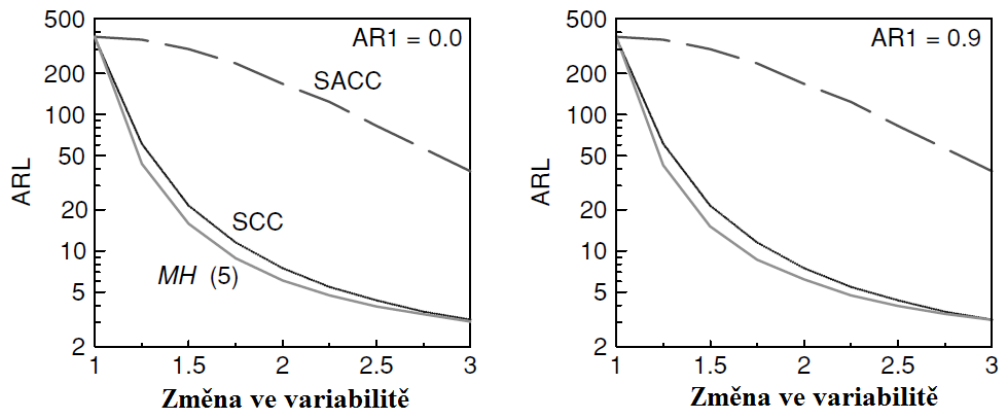
Následující obrázek porovnává výkonnost ARL u SCC, SACC, a *MH* diagramu při detekci posunu střední hodnoty procesu AR(1). Grafy ARL pro SCC graf jsou získány pomocí programu, který publikoval Wärdell et al. (1994). SACC je založen na prvních 15 autokorelacích 200 nejnovějších reziduí. Konstanta regulační meze  $D$  je nastavena na hodnotu 2,835 a  $\lambda = 1$ . Grafy ARL pro SACC jsou vypočítány s použitím simulace Monte Carlo. Výsledné ARL jsou znázorněny na následujícím obrázku. Kromě posunů střední hodnoty, které mohou produkovat velké hodnoty reziduí, bude zřejmé, že diagram *MH* je mnohem citlivější při detekci posunu střední hodnoty než SCC. Ve všech případech, diagram *MH* překonává svého hlavního soupeře, diagram SACC.



**Obr. 5.2** Chování ARL diagramů SCC, SACC a *MH* při detekci posunu procesní střední hodnoty (Zdroj: Vlastní zpracování)

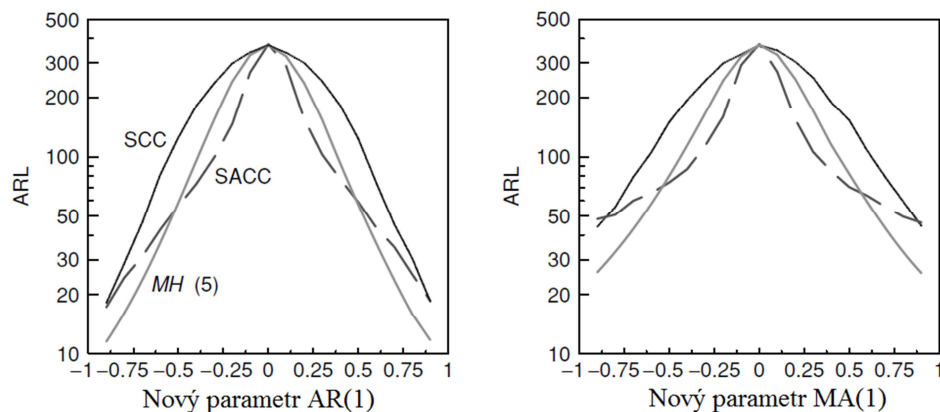
### 5.1.1.3 Citlivost na posun rozptylu

Přestože SACC může reagovat na změny v procesním rozptylu, není příliš citlivý na tyto změny (Atienza et al., 1997). Analogicky, k Shewhartovým regulačním diagramům pro jednotlivé hodnoty, které obsahují, jak informaci o střední hodnotě, tak i o rozptylu řady, SCC má lepší ARL výkon při detekci posunu ve variabilitě než SACC.



**Obr. 5.3** Citlivost diagramů SCC, SACC a MH při detekci posunu ve variabilitě  
(Zdroj: Vlastní zpracování)

Rozšířením daného porovnání tak, aby zahrnovalo navrhovaný *MH* graf, můžeme vidět z následujícího obrázku, že *MH* graf dosahuje při monitorování změn v rozptylu nejlepší výkonnosti po stránce ARL.



**Obr. 5.4** Citlivost grafů SCC, SACC a MH při detekci posunů pro různé parametry modelu (Zdroj: Vlastní zpracování)

#### 5.1.1.4 Citlivost na detekci změn v autokorelační struktuře

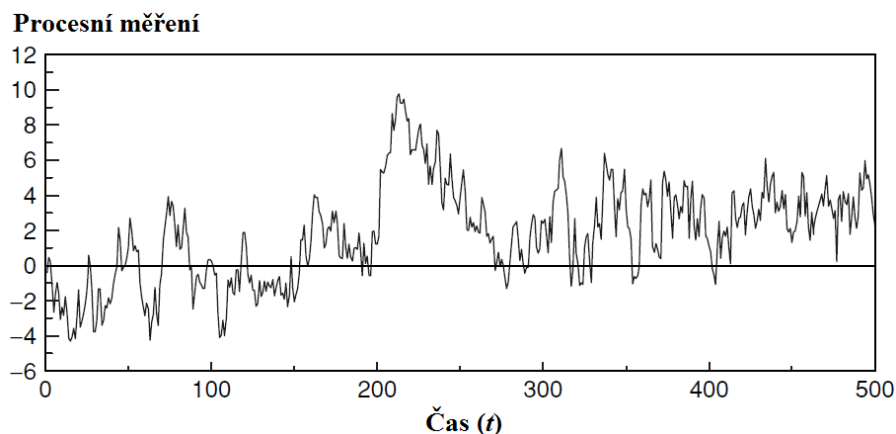
Dalším důležitým ukazatelem přítomnosti vnější příčiny ovlivnění procesu variabilitou je změna v autokorelační struktuře  $\{x_t\}$ . Za předpokladu, že proces začíná v režimu pod kontrolou, tedy splňuje podmínky i.i.d. a  $N(0,1)$ , předchozí obrázek porovnává citlivost SCC, SACC a *MH* grafu při detekci změn v autokorelační struktuře dané řady. Je zřejmé, že SACC je lepší při detekci malých posunů v AR(1), nebo MA(1) parametru. Pro velké posuny, je *MH* diagram mnohem citlivější. Tyto výsledky jsou očekávány, jelikož SACC

je založen na velkém počtu výběrů, zatímco diagram  $MH$ , který je analyzován na předchozím obrázku 5.4, je založen pouze na  $n = 5$ .  $MH$  diagram je také lepší než  $SCC$ .

### 5.1.2 Výpočetní simulační studie

Následně bude demonstrována aplikace navržené regulační techniky pomocí simulace změn v procesu  $AR(1)$ . K reprezentaci stavu procesu pod kontrolou, bude generováno 200 pozorování pocházejících z procesu  $AR(1)$  s  $\phi_1 = 0,9$ ,  $\mu = 0$  a  $\sigma_\varepsilon = 1$ . Od hodnoty  $t = 201$  byla posunuta střední hodnota o  $1,5\sigma$ . V hodnotě  $t = 301$  se střední hodnota vrátila zpět do původního stavu, tedy  $\mu = 0$ , ale rozptyl se zvýšil o 20 %. V  $t = 401$ , se chyba rozptylu vrátila na původní úroveň, ale proces se změnil na model  $ARMA(1,1)$  s  $\phi_1 = 0,9$  a  $\theta = 0,6$ . Výsledná řada je znázorněna na následujícím obrázku.

Pro ilustraci techniky výpočtu, byly tyto statistiky vyneseny v  $MH$  grafu, procesní pozorování pro  $t = 196$  do  $t = 205$  jsou uvedeny v následující tabulce. Při  $n = 5$ , regulační mez, která bude produkovat průměrnou délku kroku pod kontrolou 370, tedy  $ARL = (1/0,0027 = 370)$ , je přibližně 16,32.



**Obr. 5.5** Simulovaná měření daného procesu (Zdroj: Vlastní zpracování)

V  $t = 200$  musí být  $MH$  statistika založena na vektoru dat  $\mathbf{x}_{200}^* = (x_{196}, x_{197}, \dots, x_{200})'$ . Pro stav pod kontrolou, odpovídající kovarianční a její inverzní matice jsou následující:

$$\Sigma_0 = \begin{bmatrix} 5.263 & 4.737 & 4.263 & 3.837 & 3.453 \\ 4.737 & 5.263 & 4.737 & 4.263 & 3.837 \\ 4.263 & 4.737 & 5.263 & 4.737 & 4.263 \\ 3.837 & 4.263 & 4.737 & 5.263 & 4.737 \\ 3.453 & 3.837 & 4.263 & 4.737 & 5.263 \end{bmatrix} \text{ a}$$

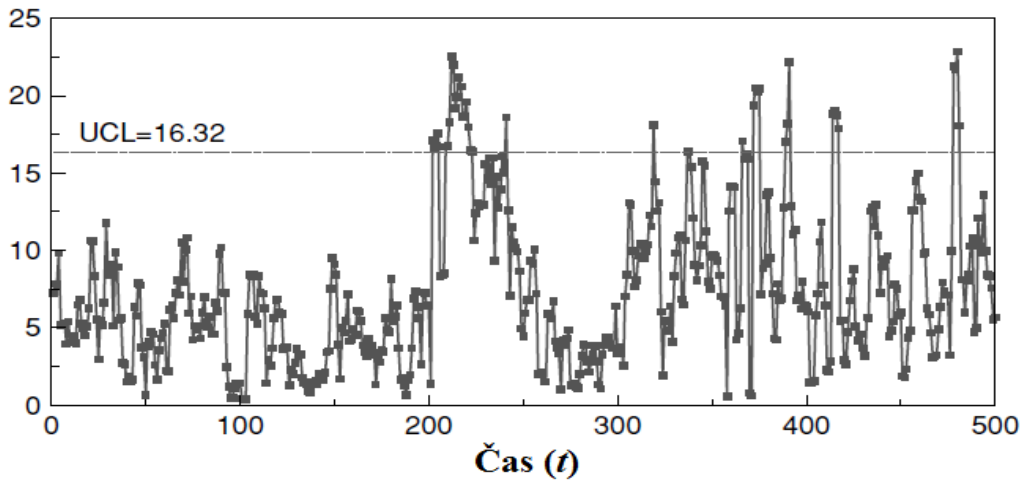
$$\Sigma_0^{-1} = \begin{bmatrix} 1.00 & -0.90 & 0.00 & 0.00 & 0.00 \\ -0.90 & 1.81 & -0.90 & 0.00 & 0.00 \\ 0.00 & -0.90 & 1.81 & -0.90 & 0.00 \\ 0.00 & 0.00 & -0.90 & 1.81 & -0.90 \\ 0.00 & 0.00 & 0.00 & -0.90 & 1.81 \end{bmatrix} .$$

**Tab. 5.1** Simulovaný proces měření,  $196 \leq t \leq 205$  (Zdroj: Vlastní zpracování)

$t$	$x_t$	Statistika $MH$	$t$	$x_t$	Statistika $MH$
196	-0,577	2,644*	201	1,666	1,383
197	1,934	7,311*	202	5,470	17,128
198	1,968	7,226*	203	5,315	16,546
199	1,234	6,426*	204	5,265	16,761
200	1,224	6,437	205	5,672	17,555

\* Vypočteno z pozorování, která nejsou znázorněna v této tabulce

**Statistika  $MH$**



**Obr. 5.6** Diagram  $MH$  pro řadu z předchozího obrázku

(Zdroj: Vlastní zpracování)

V časovém okamžiku  $t = 200$ , je statistika  $MH$  proto dána

$$MH = [-0.577 \ 1.934 \ 1.968 \ 1.234 \ 1.224] \begin{bmatrix} 1.00 & -0.90 & 0.00 & 0.00 & 0.00 \\ -0.90 & 1.81 & -0.90 & 0.00 & 0.00 \\ 0.00 & -0.90 & 1.81 & -0.90 & 0.00 \\ 0.00 & 0.00 & -0.90 & 1.81 & -0.90 \\ 0.00 & 0.00 & 0.00 & -0.90 & 1.00 \end{bmatrix} \begin{bmatrix} -0.577 \\ 1.934 \\ 1.968 \\ 1.234 \\ 1.224 \end{bmatrix} =$$

$$= 6.437$$

Výše uvedené hodnoty mohou být snadno vypočítány pomocí standardního tabulkového procesoru, jako je Excel nebo Lotus1-2-3. Další statistika  $MH$ , která je založena na vektoru pozorování  $\mathbf{x}_{201}^* = (x_{197}, x_{198}, \dots, x_{201})'$ , mohou být vypočítány stejným způsobem. Statistika  $MH$  pro  $t = 196$  do  $t = 205$  jsou uvedeny v předchozí tabulce. Celý regulační diagram pro řadu z obr. 5.5 je znázorněn na předchozím obrázku 5.6. Je zřejmé, že navržený regulační diagram účinně detekuje simulované procesní odchylky.

Většina moderních průmyslových procesů produkují měření, která jsou autokorelována. Chceme-li efektivně detekovat přítomnost procesních odchylek způsobených vnějšími zdroji variability, musíme současně sledovat změny v MVAS těchto procesních měření. Jedním ze způsobů, jak monitorovat autokorelaci procesu je používat tři regulační diagramy (jeden pro detekci změny střední hodnoty, druhý pro rozptyl, a třetí pro autokorelační strukturu dané řady). Jinou alternativou je vytvořit monitorovací proceduru, která může současně detekovat změny v MVAS. Pomocí vlastností stacionárního Gaussovského ARMA procesu, byl vyvinut regulační aparát založený na statistice  $H$  (5.9). Ve srovnání s ostatními Shewhartovými aparáty pro monitorování autokorelovaných procesů, je navrhovaný klouzavý diagram  $H$  účinný v simultánní detekci posunů v MVAS dané řady.

## 5.2 Návrh aparátu pro monitorování autokorelovaných procesů

V této části svojí diplomové práce nejprve poskytnu stručný přehled vědeckých prací na téma statistického řízení procesů pro autokorelované procesy. Toto téma je v současné době velmi frekventované, jelikož ve výrobních podnicích, při automatickém sběru dat a následném vyhodnocování se neustále setkáváme s autokorelací v časových řadách. Dále pak na simulovaných procesních datech budu aplikovat přístup navržený Tangem a Atienzou (2002), který zahrnuje monitorování určité statistiky, využívající metodiku modelování časových řad. Výkonnost tohoto navrženého monitorovacího aparátu

vyhodnotím a následně jej porovnam s tradičními metodami založenými na monitorování reziduí modelu časové řady.

Klasické metody statistického řízení procesů (SPC), jako jsou Shewhartovy regulační diagramy a CUSUM regulační diagramy předpokládají, že získaná procesní data jsou nezávislá. Nicméně, tento předpoklad byl zpochybněn, jakmile bylo zjištěno, že v mnoha praktických situacích, jsou data sériově korelovány. Výkonnost klasických regulačních diagramů se výrazně zhorší v případě autokorelace v datech. To motivovalo průkopnickou práci Alwana a Robertse (1988), kteří navrhli monitorování očekávaných chyb po odhadu vhodného modelu časové řady daného procesu. Tato metoda je intuitivní, jakmile je autokorelace vysvětlena základním modelem časové řady, a zároveň rezidua splňují podmínky nezávislých náhodných chyb procesu. Tradiční SPC metody mohou tedy být použity k monitorování reziduí. Vědecké práce na tento problém je možné rozdělit zhruba do dvou témat; na ty, které jsou založeny na modelech časových řad a ty, které jsou na modelu nezávislé. Pokud jde o první, byly navrženy již tři obecné přístupy: ty, které monitorují rezidua, ty založené na přímých pozorování a ty založené na nových statistických charakteristikách. Stručný přehled těchto přístupů je uveden v dalším textu. Wardell et al. (1994) a Lu a Reynolds (1999) navrhli pro monitorování reziduí použití regulačního diagramu exponenciálních vážených klouzavých průměrů (EWMA). Apley a Shi (1999) navrhli přístup na základě všeobecného věrohodnostního poměrového testu (GLRT) při zjišťování posunu střední hodnoty v autokorelovaných procesech. Apley a Tsung (2002) navrhli diagram aktivačního kumulativního skóre (cuscore), který má podobný charakter jako GLRT, ale je jednodušší k implementaci. Castagliola a Tsung (2005) zkoumali vliv nenormality na reziduální regulační diagramy. Navrhli upravený Shewhartův regulační diagram pro rezidua – diagram zvláštních příčin (SCC), který je více robustní v podmínkách nenormality. Testik (2005) považoval za nejistotu v modelu časové řady odhad parametrů a navrhl řešení v podobě širších EWMA kontrolních mezí pro monitorování reziduí autoregresního procesu prvního řádu AR (1).

Vezměme nyní v úvahu jiný přístup zahrnující přímé monitorování procesních výstupů namísto reziduí. Montgomery a Mastrangelo (1991) navrhli klouzavý centrovaný EWMA (MCEWMA) diagram, který aproximuje model časové řady EWMA modelem. Mastrangelo a Brown (2000) poskytli další výzkum vlastností diagramu MCEWMA. Alwan (1992) vytvořil obecnou strategii ke studiu účinků autokorelace pomocí Shewhartových regulačních diagramů. Timmer et al. (1998) porovnali různé regulační



diagramy CUSUM pro AR (1) procesy. Lu a Reynolds (1999) zkoumali výkon regulačních diagramů CUSUM jak pro rezidua, tak pro pozorování, a dospěli k závěru, že jsou téměř stejně účinné. Tito autoři tvrdí, že výkonnost diagramů CUSUM a EWMA při odhalování posunu střední hodnoty je srovnatelná. Atienza a kol. (2002) také navrhli metodu CUSUM pro autokorelovaná pozorování. Dále také navrhli novou testovací statistiku pro detekci aditivních odlehlých hodnot ( $AO^2$ ), inovačních odlehlých hodnot ( $IO^3$ ) a posunu úrovně ( $LS^4$ ). Modely časových řad založené na tomto přístupu jsou snadné na pochopení a v některých situacích jsou velmi účinné. Nicméně, tento přístup vyžaduje identifikaci vhodného modelu časové řady z množiny počátečních dat, která jsou pod kontrolou. V praxi nemusí být snadné tento model zjistit, a pro inženýry kvality se může zdát být příliš složité tuto techniku realizovat. Proto přístup nezávislý na modelu v poslední době přitahuje velkou pozornost. Nejoblíbenějším přístupem nezávislým na modelu je vytvoření vícerozměrné statistiky pro autokorelovaný jednorozměrný proces (Fuchs a Kenett, 1998), a poté jeho monitorování s odpovídajícím vícerozměrným regulačním diagramem. Krieger et al. (1992) použili vícerozměrnou metodu CUSUM. Apley a Tsung (2002) přizpůsobili  $T^2$  regulační diagram pro monitorování jednorozměrných autokorelovaných procesů. Atienza et al. (2002) navrhli vícerozměrný krabičkový (boxplot)- $T^2$  regulační diagram. Dyer et al. (2003) přizpůsobili použití vícerozměrného regulačního diagramu EWMA pro autokorelované procesy.

Další přístup nezávislý na modelu je použití regulačního diagramu dávkových průměrů, navržený Rungerem a Willemainem (1996), velmi často citovaný Montgomerym (2009) a podrobně projednáván Sunem a Xuem (2004). Hlavní výhodou tohoto přístupu spočívá v jeho jednoduchosti. Ve snaze "uvolnit" závislost, jednoduše rozdělí sekvenci pozorování do několika dávek, a pak se monitorují střední hodnoty těchto dávek na standardních regulačních diagramech pro individuální hodnoty. Další práce na téma použití SPC u autokorelovaných procesů obsahuje práce Balkina a Lina (2001), kteří zkoumali

---

<sup>2</sup> *Aditivní odlehlé hodnoty* jsou typem odlehlých hodnot, které ovlivňují jediné pozorování. Po této disturbanci se řada vrací ke své normální cestě, jako by se nic nestalo. Chyba při kódování dat může být označena jako aditivní odlehlá hodnota.

<sup>3</sup> *Inovační odlehlé hodnoty* – odlehlá hodnota, která je přírůstkem šumové složky v konkrétním bodě řady. U stacionárních řad inovační odlehlá hodnota ovlivňuje několik pozorování, u nestacionárních řad může ovlivňovat každé pozorování od určitého počátečního bodu řady.

<sup>4</sup> *Odlehlé hodnoty posunu úrovně* – posun úrovně je jako skoková funkce. Odlehlá hodnota, která od určité hodnoty posouvá všechna pozorování v řadě o jistou konstantu. Pro stacionární proces znamená úroveň posunu změnu ve střední hodnotě procesu po určitém bodě, kdy je proces transformován na nestacionární.

používání citlivostních pravidel pro Shewhartovy regulační diagramy při autokorelovaných datech. Dále Zhang (2000), Noorossana a Vaghefi (2005), Kalgonda a Kulkarni (2004), kteří nahlédli do regulačních diagramů pro vícerozměrné autoregresní procesy. Ačkoli bylo napsáno spousta prací na téma statistického řízení autokorelovaných procesů, doposud neexistuje jediná metoda, která je všeobecně aplikovatelná, nebo se ukázala být účinná ve většině situací.

Na tomto místě se domnívám, že nejslibnějším přístupem je použití nově koncipované statistiky, kterou následně představím. V následujícím textu uvedu aparát, navržený Atienzou, pro monitorování autokorelovaných procesů. Zaměřím se na studium chování statistiky  $\lambda_{LS,t}$  používanou pro detekci úrovně posunu v procesu AR(1). Výkonnost tohoto aparátu bude vyhodnocena a srovnána s Alvanovým a Robertsovým aparátem SCC (diagram zvláštních příčin – pojednáno výše).

### 5.3 Zjišťování odlehlých hodnot a úrovně posunu

Budeme uvažovat model ARMA

$$\phi(B)Z_t = \phi_0 + \theta(B)\varepsilon_t, \quad (5.11)$$

kde  $Z_t$  je stacionární časová řada reprezentující proces měření,

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$$

je autoregresní polynom řádu  $p$ ,

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$$

polynom klouzavého průměru řádu  $q$ ,  $B$  je operátor zpětného posunu a  $\{\varepsilon_t\}$  je posloupnost nezávislých a normálně rozdělených náhodných chyb s nulovou střední hodnotou a konstantním rozptylem  $\sigma^2$ . Bez újmy na obecnosti budeme předpokládat, že úroveň  $\phi_0$

časové řady  $\{Z_t\}$  je nulová. Necht'  $\hat{Z}_t$  reprezentuje predikovanou hodnotu získanou z vhodně zvoleného a odhadnutého ARMA modelu, potom se rezidua

$$\{e_1 = Z_1 - \hat{Z}_1, e_2 = Z_2 - \hat{Z}_2, \dots, e_t = Z_t - \hat{Z}_t, \dots\}$$

budou chovat jako nezávislé náhodné proměnné, pocházející ze stejného rozdělení (i.i.d.).

Necht'

$$Y_t = f(t) + \frac{\theta(B)}{\phi(B)} \varepsilon_t, \quad (5.12)$$

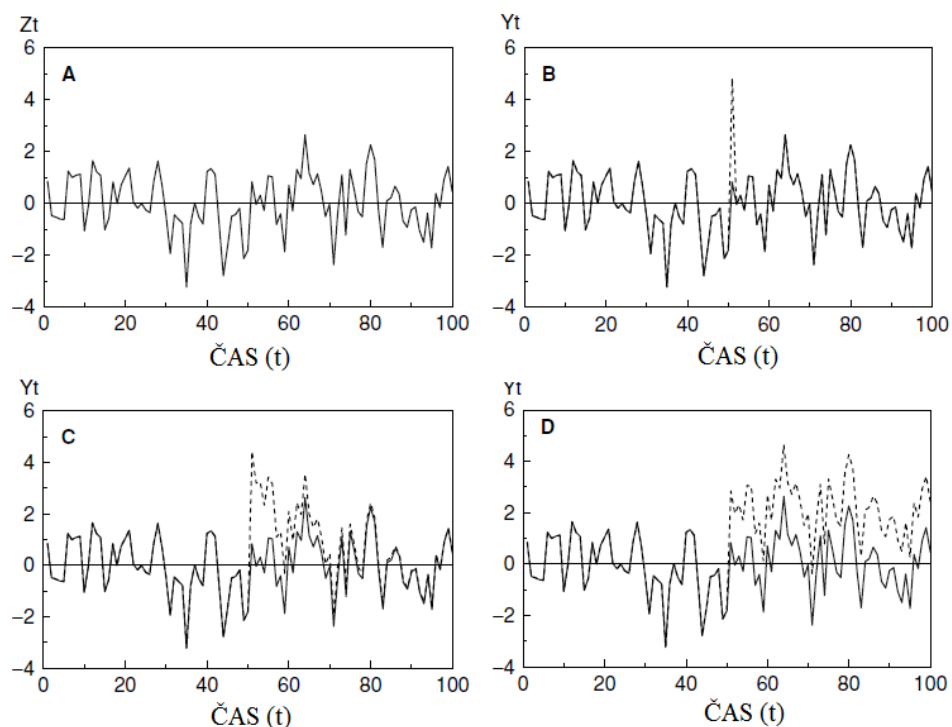
kde  $Y_t$  a  $f(t)$  reprezentují „znečištěnou“ řadu a neobvyklé exogenní disturbance, stejně jako odlehle hodnoty a různé úrovně posunu časové řady. Funkce  $f(t)$  může být deterministická nebo stochastická, závisující na typu disturbance. Pro deterministický model má  $f(t)$  následující formu

$$f(t) = \omega_0 \frac{\omega(B)}{\delta(B)} \xi_t^{(d)}, \quad (5.13)$$

kde

$$\xi_t^{(d)} = \begin{cases} 1, & \text{pokud } t = d, \\ 0, & \text{pokud } t \neq d, \end{cases}$$

je ukazatelem proměnné pro výskyt disturbance v čase  $d$ ,  $\omega(B)$  a  $\delta(B)$  jsou polynomy zpětného posunu popisující dynamický efekt disturbance na  $Y_t$ , a  $\omega_0$  je konstanta označující počáteční vliv disturbance.



**Obr. 5.7** (A) Nasimulovaná řada  $AR(1)$  s  $\phi = 0,5$ . (B) Nasimulovaná řada  $AR(1)$  s  $AO$  v  $t = 51$ . (C) Nasimulovaná řada  $AR(1)$  s  $IO$  v  $t = 51$ . (D) Nasimulovaná řada  $AR(1)$  s  $LS$  v  $t = 51$ . (Zdroj: Vlastní zpracování v jazyku R)

Pokud  $\omega(B)/\delta(B) = 1$ , disturbance má podobu aditivní odlehle hodnoty. AO tedy ovlivňuje úroveň naměřené časové řady pouze v čase  $t = d$ . Častou příčinou aditivních odlehle hodnot je chyba záznamu dat. V diskrétních výrobních procesech mohou AO nastat tehdy, pokud se v datech vyskytuje smíšená informace z velkého množství variantních surovin. Pokud  $\omega(B)/\delta(B) = \theta(B)/\phi(B)$ , rovnice (5.13) reprezentuje inovační odlehle hodnoty. AO ovlivňuje úroveň  $Y_t$  v  $t = d$ . Po dosažení  $t = d$ , tento se tento efekt ztrácí exponenciálně. IO je s největší pravděpodobností způsobena kontaminací v nepřetržitém chemickém procesu. Během preventivní údržby v chemické továrně, například v případě, pokud stávající náhradní díl (např. potrubí nebo jiné připojení) je nahrazen kontaminovanou částí, budou charakteristiky chemicky zpracovávaného procesu hrubě ovlivněny v čase  $t = d$ , ale potom efekt kontaminace bude slábnout. Pokud  $\omega(B)/\delta(B) = 1/(1 - B)$ , disturbance reprezentována rovnicí (5.13) je posun úrovně. LS posune úroveň  $Y_t$  začínající v  $t = d$  směrem nahoru nebo dolů. Tento posun přetrvává po dobu  $t > d$ . LS je způsobeno především změnou v kvalitě materiálu nebo nastavení procesu. Obrázek 5.7 ukazuje účinek těchto tří typů odlehle hodnot na procesu AR(1). V následujícím textu použijí  $\omega_{AO}$ ,  $\omega_{IO}$  a  $\omega_{LS}$  k rozlišení, zda  $\omega_0$  souvisí s AO, IO nebo LS. Z rovnic (5.12) a (5.13) dostáváme následující výraz:

$$\frac{\phi(B)}{\theta(B)} Y_t = \frac{\phi(B)\omega(B)}{\theta(B)\delta(B)} \xi_t^{(d)} \omega_0 + \varepsilon_t.$$

Pokud máme  $y_t = [\phi(B)/\theta(B)] Y_t$  a  $x_t = [\phi(B)\omega(B)/\theta(B)\delta(B)] \xi_t^{(d)}$ , dostaneme

$$y_t = \omega_0 x_t + \varepsilon_t,$$

což je jednoduchý tvar lineární regrese. Takže vliv parametru  $\omega_0$  může být odhadnut pomocí

$$\hat{\omega}_0 = \frac{\sum_{t=1}^T y_t x_t}{\sum_{t=1}^T x_t^2} \text{ s rozptylem } \text{Var}(\hat{\omega}_0) = \frac{\sigma^2}{\sum_{t=1}^T x_t^2},$$

kde  $T$  reprezentuje velikost výběru. Pomocí těchto rovnic dostaneme následující odhady  $\omega_0$  pro tři typy disturbance zmíněných výše:

$$\hat{\omega}_{AO,t} = \begin{cases} \rho_{AO,t}^2 \left( y_t - \sum_{i=1}^{T-t} \pi_i y_{t+i} \right) & t = 1, 2, \dots, T-1 \\ y_t & t = T, \end{cases} \quad (5.14)$$

$$\hat{\omega}_{IO,t} = y_t \quad t = 1, 2, \dots, T, \quad (5.15)$$

$$\hat{\omega}_{LS,t} = \begin{cases} \rho_{LS,t}^2 \left( y_t - \sum_{i=1}^{T-t} \eta_i y_{t+i} \right) & t = 1, 2, \dots, T-1 \\ y_t & t = T, \end{cases} \quad (5.16)$$

kde  $\rho_{AO,t}^2 = \left( 1 + \sum_{i=1}^{T-t} \pi_i^2 \right)^{-1}$ ,  $\rho_{IO,t}^2 = 1$  a  $\rho_{LS,t}^2 = \left( 1 + \sum_{i=1}^{T-t} \eta_i^2 \right)^{-1}$ .  $\pi_i$  a  $\eta_i$  jsou koeficienty  $B^i$

v polynomech  $\pi(B) = 1 - \pi_1 B - \pi_2 B^2 - \dots = \phi(B)/\theta(B)$

a  $\eta(B) = 1 - \eta_1 B - \eta_2 B^2 - \dots = \pi(B)/(1-B)$ . Váhy  $\pi$  jsou nalezeny násobením obou stran výrazu  $\pi(B)$ , výrazem  $\theta(B)$  pro získání  $\theta(B)(1 - \pi_1 B - \pi_2 B^2 - \dots) = \phi(B)$ . Například pro proces ARMA(1,1) dostáváme

$$\begin{aligned} 1 - \phi B &= (1 - \theta B)(1 - \pi_1 B - \pi_2 B^2 - \dots) = \\ &= 1 - (\pi_1 + \theta) B - (\pi_2 - \theta \pi_1) B^2 - (\pi_3 - \theta \pi_2) B^3 - \dots \end{aligned}$$

Porovnáním koeficientů, jako jsou mocniny  $B$ , dostáváme  $\pi_1 = \phi - \theta$ ,  $\pi_2 = \theta \pi_1$  a  $\pi_j = \theta \pi_{j-1} = \theta^{j-1} \pi_1$  pro  $j > 1$ . Podobný postup může být také použit pro výpočet vah  $\eta$ . Můžeme si velmi jednoduše ověřit, že pro model ARMA(1, 1) odpovídající váhy jsou  $\eta_1 = \phi - \theta - 1$  a  $\eta_j = \eta_{j-1} + \theta^{j-1} \pi_1$  pro  $j > 1$ . Takže pro proces AR(1) bude platit  $\pi_j = \phi^j$  a  $\eta_j = \phi - 1$  pro  $j \geq 1$ .

Pokud použijeme závěry z výše uvedeného, můžeme zkonstruovat následující testovací statistiku pro testování existence AO, IO a LS v časovém bodě  $d$ :

$$\lambda_{j,d} = \frac{\hat{\omega}_{j,d}}{\left[ \text{Var}(\hat{\omega}_{j,d}) \right]^{1/2}} = \frac{\hat{\omega}_{j,d}}{\rho_{j,d} \sigma}, \quad j = \text{AO, IO, LS}. \quad (5.17)$$

Za platnosti nulové hypotézy nepřítomnosti odlehlých hodnot nebo posunu úrovně a za předpokladu, že časový okamžik  $d$  a parametry modelu ARMA v rovnici (5.11) jsou známy, statistiky  $\lambda_{AO,t}$ ,  $\lambda_{IO,t}$  a  $\lambda_{LS,t}$  mají asymptotické rozdělení jako  $N(0,1)$ . V praxi jsou

tyto parametry časové řady pro tyto statistiky obvykle známy, a proto musí být nahrazeny nějakými konzistentními odhady.

Pro detekci AO, IO nebo LS v neznámé pozici vypočteme následující statistiky:

$$\lambda_{j,\max} = \max_{1 \leq t \leq T} \{ \lambda_{j,t} \} \quad j = \text{AO, IO, LS.} \quad (5.18)$$

Nulová hypotéza nepřítomnosti AO, IO nebo LS je zamítnuta, pokud  $\lambda_{j,\max}$  překročí předem stanovenou kritickou hodnotu. Poznamenejme, že je obtížné určit přesné opakování vzorkování distribuce v rovnici (5.18). Chen a Liu (1993) uvažovali široké spektrum vah  $\pi$  vzorů v odhadu percentilů statistiky (5.18) pomocí Monte Carlo simulace. Poznamenali, že percentily související s (5.16) jsou výrazně nižší než ty, které souvisejí s rovnicemi (5.14) a (5.15). Pro proces AR(1), například, pokud  $T = 200$ , odhad 1. percentilu pro  $\lambda_{\text{LS},\max}$ , na základě závěru Chena a Liua, se pohybuje v rozmezí 3,3 a 3,5, zatímco pro  $\lambda_{\text{AO},\max}$  a  $\lambda_{\text{IO},\max}$  se pohybuje v rozmezí 4,0 a 4,2.

V SPC, se především zabýváme detekcí přítomnosti vnějších zdrojů variability projevujícími se změnami v procesní úrovni nebo rozptylu. Pokud se nám podaří zjistit, zda změna v chování procesu je způsobena díky AO, IO nebo LS, můžeme tímto zúžit její nejpravděpodobnější příčinu. To umožní rychlé odhalení a nápravu problémů spojených se zjištěným signálem o změně procesu. V následujícím textu můžeme vidět, že systém SPC na základě detekce statistik odlehlých hodnot a změny úrovně procesu může být vytvořeno pro účely monitorování procesů.

## 5.4 Vlastnosti $\lambda_{\text{LS},t}$

Je známo, že pokud dojde k odlehlé hodnotě nebo posunu úrovně v nějakém známém bodě, statistiky (5.14), (5.15) a (5.16) jsou normálně rozděleny se středními hodnotami  $\omega_{\text{AO}}$ ,  $\omega_{\text{IO}}$  a  $\omega_{\text{LS}}$  a rozptyly  $\rho_{\text{AO},d}^2 \sigma^2$ ,  $\rho_{\text{IO},d}^2 \sigma^2$  a  $\rho_{\text{LS},d}^2 \sigma^2$ . V této části se budeme zabývat vlastnostmi  $\{ \lambda_{\text{LS},t} : 1 \leq t \leq T \}$ , když dojde k posunu úrovně nebo odlehlé hodnotě. Přitom výše uvedený postup v předchozí části může vést ke způsobu, jak tento použít citlivěji pro detekci přítomnosti exogenních zásahů nebo zvláštních příčin, pokud je implementován on-line.

### 5.4.1 Efekt úrovně posunu

Při studiu efektu odlehlých hodnot a úrovně posunu na  $\{ \lambda_{\text{LS},t} : 1 \leq t \leq T \}$ , jsme se zaměřili na model AR(1)

$$Z_t - \mu = \phi(Z_{t-1} - \mu) + \varepsilon_t, \quad t = 1, 2, \dots, T, \quad (5.19)$$

kde  $-1 < \phi < 1$  a náhodné chyby  $\varepsilon_t$  jsou i.i.d.  $N(0, \sigma^2)$ . Pro proces AR(1) popsany rovnicí (5.19),  $\pi(B) = 1 - \phi B$  a  $\eta(B) = 1 - (\phi - 1)B - (\phi - 1)B^2 - \dots$ . Nyní předpokládejme, že v čase  $t = d$  ( $1 \leq t \leq T$ ) se střední hodnota posune z 0 do  $\omega_{LS}$ . Pomocí Tsayovy (1988) notace dostáváme  $Y_t = Z_t$  pro  $t < d$  a  $Y_t = Z_t + \omega_{LS}$  pro  $t \geq d$ . Potom může být jednoduše ukázáno, že

$$y_t = \begin{cases} \varepsilon_t, & t < d \\ \omega_{LS} + \varepsilon_t, & t = d \\ (1 - \phi)\omega_{LS} + \varepsilon_t & t > d. \end{cases} \quad (5.20)$$

Z (5.17) můžeme vidět, že  $\hat{\omega}_{LS,t}$  je lineární kombinací normálně rozdělených náhodných veličin, a proto je sám normálně rozdělen. Očekávanou hodnotu a rozptyl  $\hat{\omega}_{LS,t}$  pro  $1 \leq t \leq T$  můžeme odvodit následovně. Pro  $t < d$ ,

$$\begin{aligned} E(\hat{\omega}_{LS,t}) &= E\left[\rho_{LS,t}^2 \left( y_t - \sum_{i=1}^{T-t} (\phi - 1) y_{t+i} \right)\right] = \\ &= \rho_{LS,t}^2 \left[ E(y_t) + (1 - \phi) \left( \sum_{i=1}^{d-1} E(y_i) + E(y_d) + \sum_{i=d+1}^T E(y_i) \right) \right] = \\ &= \rho_{LS,t}^2 \omega_{LS} (1 - \phi) [1 + (T - d)(1 - \phi)]. \end{aligned}$$

Pro  $t = d$ ,

$$\begin{aligned} E(\hat{\omega}_{LS,t}) &= E\left[\rho_{LS,d}^2 \left( y_d - \sum_{i=1}^{T-d} (\phi - 1) y_{d+i} \right)\right] = \\ &= \rho_{LS,d}^2 \left[ E(y_d) + (1 - \phi) \sum_{i=d+1}^T E(y_i) \right] = \\ &= \omega_{LS}. \end{aligned}$$

A konečně pro  $t > d$ ,

$$\begin{aligned} E(\hat{\omega}_{LS,t}) &= E\left[\rho_{LS,t}^2 \left( y_t - \sum_{i=1}^{T-t} (\phi - 1) y_{t+i} \right)\right] = \\ &= \rho_{LS,t}^2 \left[ E(y_t) + (1 - \phi) \sum_{i=t+1}^T E(y_i) \right] = \\ &= \rho_{LS,t}^2 \omega_{LS} (1 - \phi) [1 + (T - t)(1 - \phi)]. \end{aligned}$$

Dále můžeme ukázat, že rozptyl  $\hat{\omega}_{LS,t}$  pro  $1 \leq t \leq T$  je  $\rho_{LS,t}^2 \sigma^2$ . Takže  $\{\lambda_{LS,t}; 1 \leq t \leq T\}$  má následující očekávanou vlastnost:

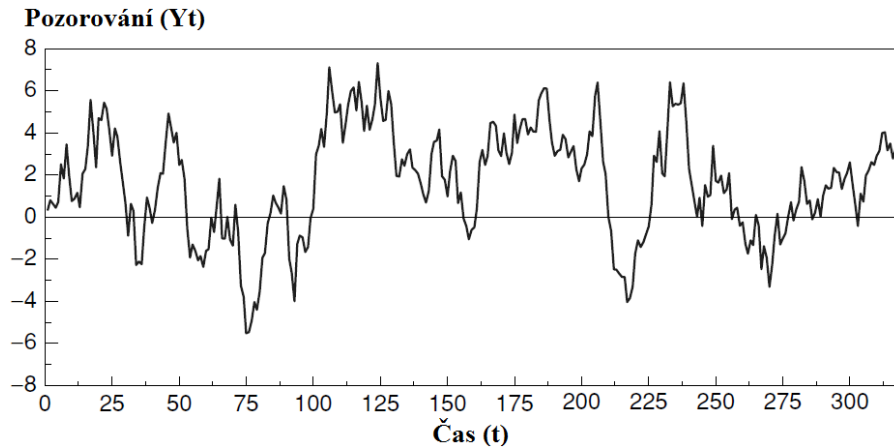
$$E[\lambda_{LS,t}] = \begin{cases} \rho_{LS,t} \omega_{LS} (1-\phi) [1+(T-d)(1-\phi)] / \sigma, & 1 \leq t < d, \\ \omega_{LS} / \rho_{LS,t} \sigma, & t = d, \\ \rho_{LS,t} \omega_{LS} (1-\phi) [1+(T-t)(1-\phi)] / \sigma, & d < t \leq T. \end{cases} \quad (5.21)$$

Z (5.21) můžeme vidět, že maximum  $E[\lambda_{LS,t}]$  je dosaženo v  $t = d$ . Což odpovídá Tsayově statistice pro detekci úrovně posunu. To může být patrné z (5.21), že když se dosáhne očekávané hodnoty  $\lambda_{LS,t}$  stává se nenulová pro  $t \neq d$ . Takže kromě velikosti  $\lambda_{LS,t}$ , můžeme také zkoumat informaci danou  $\lambda_{LS,d}$  pro  $t \neq d$  při odhalování LS.

K ilustraci výše zmíněných závěrů, nasimuluji proces AR(1) s  $\phi = 0,9$  a  $\sigma = 1$ . Pro  $t = 1, \dots, 100$  je střední hodnota časové řady 0. Počínajíc v čase  $t = 101$ , zavedu LS o velikosti  $\sigma_z = 2,254157$  (viz obrázek 5.8). Pomocí SCC (Shewhartův regulační diagram) s regulačními mezemi stanovenými na úrovni  $\pm 3\sigma$ , můžeme vidět, že odpovídající ARL je 223,30 se směrodatnou odchylkou 283,60. Pro data z obrázku 5.8, změna ve střední hodnotě byla odhadnuta pomocí SCC v  $t = 318$  nebo 218 hodnot po realizaci změny (viz obrázek 5.9). Tato skutečnost ukazuje na nízkou citlivost SCC. Pro ilustraci, jak průměr statistiky  $\lambda_{LS,t}$  pro  $t = 1 \dots, T$  může být použit pro detekci změny úrovně časové řady, bude spočteno a vykresleno  $\{\lambda_{LS,t}; 1 \leq t \leq T\}$  před a po změně, která nastane v  $T = 100$ . Výpočty uvedeny v tabulce 5.2 byly provedeny pomocí nástroje MS Excel 2010. V tabulce,  $y_t$  se předpokládalo rovno nule, což je typické v prediktivních situacích. Nicméně, v on-line monitorování procesů, jak je popsáno v následujícím textu, je možné stanovit hodnotu  $y_1$  na základě předchozí sady procesních měření. Jelikož pro proces AR(1)  $\eta_j = \phi - 1 (j \geq 1)$ ,

výraz  $\sum_{i=1}^{T-t} \eta_i y_{t+i}$  v rovnici (5.16) můžeme snadno vypočítat, když vynásobíme každé  $y_t$  s  $\phi - 1$  (tj. 4. sloupec) a následně vypočítáme dílčí součty ve sloupci 5. Když víme, že  $\rho_{LS,t}^2 = [1+(T-t)(\phi-1)]^{-1}$ , můžeme nyní pomocí hodnot ve sloupcích 3 a 5 vypočítat  $\lambda_{LS,t}$ . Data ve sloupcích 5 a 6 jsou spočteny pro  $T = 100$ , zatímco hodnoty ve sloupcích 7 a 8 jsou pro  $T = 106$ .





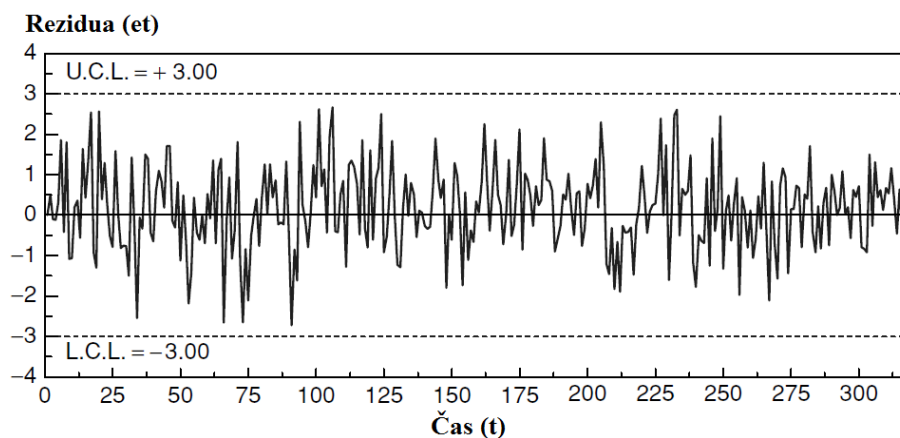
**Obr. 5.8** Simulovaná časová řada  $AR(1)$  ( $\phi = 0,9$ ,  $\sigma^2 = 1,0$  a  $\omega_{LS,101} = \sigma_z$ )

(Zdroj: Vlastní zpracování v jazyku R)

V programu R, jsem využil následující výpis:

```
ar1.model = list(ar=0.9)
> mu = 0
> set.seed(123)
> ar1.sim = mu + arima.sim(model=ar1.model,n=325)
> ar1.acf = ARMAacf(ar=0.9, ma=0, lag.max=10)
> par(mfrow=c(2,1))
> ts.plot(ar1.sim,main="AR(1) Process: mu=0, phi=0.9",
+ xlab="time",ylab="y(t)", col="black", lwd=2)
```

Obrázek 5.10 ukazuje, že když  $T = 100$  (k žádné změně nedošlo), průměrná hodnota  $\{\lambda_{LS,t} : 1 \leq t \leq T\}$  je přibližně nula. Ze šesti pozorování po zavedení LS v  $t = 101$  můžeme snadno vidět, že se posunula průměrná hodnota  $\lambda_{LS,t}$ . To právě ukazuje citlivost průměru  $\lambda_{LS,t}$  při detekci přítomnosti LS.



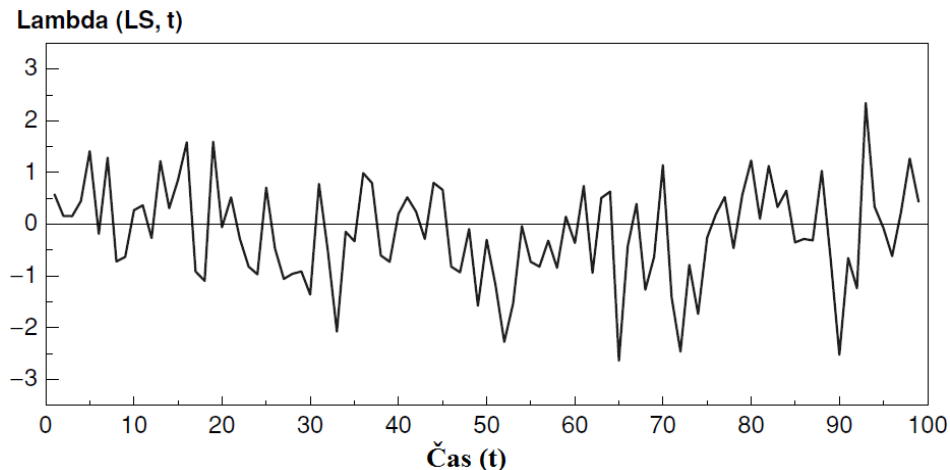
**Obr. 5.9** Diagram zvláštních příčin pro řady z obrázku 5.8

(Zdroj: Vlastní zpracování v jazyku R)

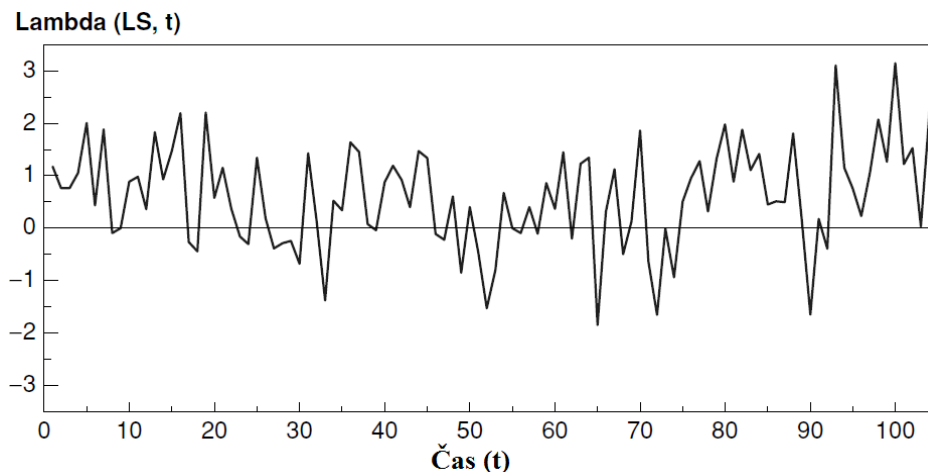
Tab. 5.2 Výpočet charakteristiky  $\lambda_{LS,t}$  pomocí MS Excel 2010

(Zdroj: Vlastní zpracování)

$t$	$Y_t$	$y_t$	$ny_t$	$\eta \sum_{i=1}^{100-t} y_{t+i}$	$\lambda_{LS,t}(T = 100)$	$\eta \sum_{i=1}^{106-t} y_{t+i}$	$\lambda_{LS,t}(T = 106)$
1	0,3434	0,0000	0,0000	-0,3669	0,2601	-1,2310	0,8598
2	0,8053	0,4962	-0,4960	-0,3173	0,5781	-1,1814	1,1746
3	0,6245	-0,1003	0,0100	-0,3273	0,1618	-1,1914	0,7658
4	0,4441	-0,1179	0,0118	-0,3391	0,1580	-1,2032	0,7636
5	0,7227	0,3230	-0,0323	-0,3068	0,4511	-1,1709	1,0538
6	2,5025	1,8520	-0,1852	-0,1216	1,4170	-0,9857	2,0066
7	1,8366	-0,4157	0,0416	-0,1632	-0,1818	-1,0273	0,4335
8	3,4523	1,7994	-0,1799	0,0167	1,2865	-0,8473	1,8809
9	2,0216	-1,0854	0,1085	-0,0918	-0,7190	-0,9559	-0,0923
10	0,7554	-1,0640	0,1064	-0,1982	-0,6281	-1,0623	-0,0012
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
90	0,8240	-0,5045	0,0504	0,1827	-0,6552	-0,6814	0,1643
91	-1,9793	-2,7209	0,2721	-0,0894	-2,5205	-0,9535	-1,6481
92	-2,6344	-0,8530	0,0853	-0,1747	-0,6527	-1,0388	0,1740
93	-3,9863	-1,6153	0,1615	-0,3363	-1,2365	-1,2004	-0,3903
94	-1,2763	2,3113	-0,2311	-0,1051	2,3471	-0,9692	3,0998
95	-0,8799	0,2688	-0,0269	-0,0783	0,3387	-0,9423	1,1495
96	-0,9547	-0,1628	0,0163	-0,0945	-0,0670	-0,9586	0,7587
97	-1,6546	-0,7953	0,0795	-0,1741	-0,6121	-1,0382	0,2326
98	-1,4208	0,0683	-0,0068	-0,1672	0,2332	-1,0313	1,0581
99	-0,0450	1,2337	-0,1234	-0,0439	1,2712	-0,9080	2,0704
100	0,3982	0,4387	-0,0439			-0,8641	1,2653
101	2,9764	2,6181	-0,2168			-0,6023	3,1428
102	3,3955	0,7167	-0,0717			-0,5306	1,2231
103	4,1854	1,1294	-0,1129			-0,4177	1,5244
104	3,3373	-0,4295	0,0430			-0,4606	0,0308
105	4,9489	1,9453	-0,1945			-0,2661	2,2004
106	7,1148	2,6609	-0,2661				



**Obr. 5.10** Vlastnosti a chování  $\lambda_{LS,t}$  pro  $t = 1, \dots, T = 100$  (bez posunu)  
(Zdroj: Vlastní zpracování v jazyku R)



**Obr. 5.11** Vlastnosti a chování  $\lambda_{LS,t}$  pro  $t = 1, \dots, T = 100$  (úroveň posunu velikosti  $\omega_{LS,101} = \sigma_z$  v  $t = 101$ ) (Zdroj: Vlastní zpracování v jazyku R)

#### 5.4.2 Vliv odlehlých hodnot

Kromě úrovně posunů, jsme zjistili, že dalším společným projevem zvláštních příčin variability je přítomnost aditivních nebo inovačních odlehlých hodnot v časové řadě. Přestože zjištění přítomnosti AO nebo IO není naším prvořadým zájmem, je zajímavé vědět, jak se bude  $\lambda_{LS,t}$  chovat, když jsou přítomny tyto dva typy odlehlých hodnot. Pomocí přístupu, který je uveden v předchozí části můžeme ukázat, že když dojde k AO v rozmezí mezi 1 a T,  $\hat{\omega}_{LS,t}$  má očekávané následující chování:

$$E(\hat{\omega}_{LS,t}) = \begin{cases} \rho_{LS,t}^2 \omega_{AO} (1-\phi)^2, & 1 \leq t < d, \\ \rho_{LS,t}^2 \omega_{AO} [1-\phi(1-\phi)], & t = d, \\ -\rho_{LS,t}^2 \omega_{AO} \phi, & t = d+1, \\ 0, & d+1 < t \leq T. \end{cases}$$

Přestože první hodnoty  $d-1\hat{\omega}_{LS,t}$  jsou ovlivněny díky  $\omega_{AO,d}$ , jejich rozsah je výrazně deflován ve srovnání s rozsahem AO, neboť obě hodnoty  $\rho_{LS,t}^2$  a  $1-\phi$  jsou pro stacionární AR(1) proces menší než 1. To je také případ pro  $t = d$ . To naznačuje, že  $\hat{\omega}_{LS,t}$  není vhodné pro zjišťování přítomnosti AO. Nicméně, je zřejmé, že když  $T = d$ ,  $E(\hat{\omega}_{LS,d}) = E(\hat{\omega}_{AO,d}) = \omega_{AO,d}$ . Takže když k AO dojde na konci časové řady, může být identifikována pomocí  $\hat{\omega}_{LS,t}$  v závislosti na velikosti  $\omega_{AO}$ . Protože se právě jedná o případ, kdy se zabýváme SPC daty, regulační diagramy založené na  $\hat{\omega}_{LS,t}$  mohou být efektivní při detekci aditivních odlehlých hodnot. Pro inovativní odlehlé hodnoty, vyskytující se v rozmezí 1 a T, má  $\lambda_{LS,t}$  následující očekávané vlastnosti:

$$E(\hat{\omega}_{LS,t}) = \begin{cases} \rho_{LS,t}^2 \omega_{IO} (1-\phi), & 1 \leq t < d, \\ \rho_{LS,t}^2 \omega_{IO}, & t = d, \\ 0, & d < t \leq T. \end{cases}$$

Vlastnosti  $\hat{\omega}_{LS,t}$  v případě, kdy dojde k IO, jsou v podstatě stejné, jako v případě, když dojde k AO. To znamená, že  $\hat{\omega}_{LS,t}$  nemůže vhodně zpracovat IO, když nastane v rozmezí 1 až T. Podobně jako v případě AO, když dojde k IO v  $T = d$ ,  $E(\hat{\omega}_{LS,d}) = E(\hat{\omega}_{IO,d}) = \omega_{IO,d}$ . Takže SPC aparát, založený na  $\hat{\omega}_{LS,t}$  mohou rovněž identifikovat přítomnost inovativních odlehlých hodnot.

## 5.5 Návrh monitorovacího aparátu

Výsledky v předchozích částech naznačují, že můžeme vytvořit regulační proceduru pro zjišťování úrovně procesního posunu pomocí monitorování buď  $\lambda_{LS,\max}$  nebo

$\bar{\lambda}_{LS,t} = \sum_{i=1}^T \lambda_{LS,i} / T$ . Při posuzování vlastností navrhovaných regulačních diagramů, se

zaměříme na zjištění LS v procesu AR (1).

### 5.5.1 Implementace regulačního diagramu

Při implementaci navrženého regulačního aparátu je třeba začít s počáteční množinou  $m$  pozorování při nastavení procesu, který je pod kontrolou (ve statisticky zvládnutém stavu). V počáteční fázi implementace, je proces s největší pravděpodobností ovlivněn odlehlými hodnotami a posuny úrovně. Jejich přítomnost může vést ke špatné identifikaci modelu. To může následně vést k chybným určením procesních odchylek. Takže v této fázi se doporučuje použití postupů, kterými se může společně dosáhnout jak detekce odlehlých hodnot, tak i odhadu modelu.

Analýza ukazuje, že regulační diagramy založené na statistice  $\lambda_{LS,t}$  jsou citlivé na volbě  $m$ . Například, výsledky simulace ukazují, že při  $m = 100$  dávají špatnou informaci o průměrné délce kroku (ARL<sup>5</sup>). Uspokojivé ARL je dosaženo při  $m = 200$ . Všimněte si, že i při velkém  $m$  dle běžných standardů SPC, to může v reálném čase monitorování procesů odpovídat jen několika vteřinám nebo minutám.

Jakmile máme k dispozici počáteční soubor  $m$  procesních pozorování, můžeme začít s výpočtem  $\lambda_{LS,t}$ . Při použití navrhovaného regulačního aparátu potřebujeme pouze zachovat  $m$  nejnovějších pozorování. To znamená, že když je k dispozici nové pozorování, odstraní se to nejstarší. Pomocí  $m$  nejnovějších pozorování, vypočteme  $m$  hodnot  $\lambda_{LS,t}$ . Pro ilustraci předpokládejme, že začneme s počátečním souborem procesních pozorování  $\{Z_1, Z_2, \dots, Z_m\}$ . Z tohoto počátečního souboru vypočítáme  $m$  hodnot  $\lambda_{LS,t}$ ,  $\{\lambda_{LS,1}, \lambda_{LS,2}, \dots, \lambda_{LS,m}\}$ . Když máme k dispozici pozorování  $Z_{m+1}$ , použijeme škálu hodnot  $\{Z_2, Z_3, \dots, Z_{m+1}\}$  k výpočtu hodnot  $\{\lambda_{LS,2}, \lambda_{LS,3}, \dots, \lambda_{LS,m+1}\}$ . Takže pro libovolné časové období  $i$  po pořízeném počátečním souboru pozorování  $m$ , máme  $\{Z_{i+1}, Z_{i+2}, \dots, Z_{m+i}\}$ , škálu  $m$  nejpozdějších pozorování, ze kterých vypočítáme  $\{\lambda_{LS,i+1}, \lambda_{LS,i+2}, \dots, \lambda_{LS,m+i}\}$ . Vyjádřeno v matematické podobě, potřebujeme sledovat buď

$$\lambda_{LS,\max,i} = \max_{i+1 \leq t \leq m+i} \{\lambda_{LS,t}\}, \quad i = 1, 2, 3, \dots,$$

nebo

---

<sup>5</sup> Kritérium **ARL** (Average Run Length) – průměrná délka kroku: udává průměrný počet bodů, než některý padne mimo regulační meze nebo dojde k rozpoznání změny (posunu) CL.

$$\bar{\lambda}_{LS,i} = \frac{\sum_{t=i+1}^{m+i} \lambda_{LS,t}}{m}, \quad i = 1, 2, 3, \dots$$

Graf signalizuje situaci statisticky nezvládnutého procesu, kdy  $\lambda_{LS,\max}$  nebo  $\bar{\lambda}_{LS,t}$  překročí předem určenou regulační mez. Tyto kontrolní meze mohou být stanoveny na základě vzorkování rozdělení  $\lambda_{LS,\max}$  nebo  $\bar{\lambda}_{LS,t}$ , což je obtížné získat, protože tyto statistiky jsou funkce závisle proměnných. Pro stanovení vhodných regulačních mezí se proto doporučuje používat simulace. To může být provedeno nejprve zadáním standardní nebo přijatelné hodnoty ARL, za předpokladu statisticky zvládnutého stavu. K určení, jaké kontrolní meze mohou být použity k dosažení této ARL hodnoty ve statisticky zvládnutém stavu, se pak používá simulační program. Pokud jde o diagram  $\lambda_{LS,\max}$ , mohou být výchozí body pro simulaci použity závěry Chena a Liua (1993).

### 5.5.2 Výkonnost, vlastnosti a chování regulačního diagramu

Pro proces AR(1) s  $\phi > 0$ , když dojde k úrovni posunutí o rozsahu  $\omega_{LS}$  v  $t = d$ , je očekávaná hodnota rezidua v daném okamžiku je  $\omega_{LS}$ . Pro  $t > d$  se očekávaná hodnota reziduí stává  $(1 - \phi)\omega_{LS}$ . Takže, v případě SCC je vysoká pravděpodobnost detekce posunu ihned, když k němu dojde. Po tomto bodě, je pravděpodobnost, že SCC detekuje tento posun, velmi malá a to zejména při  $\phi \rightarrow 1$ . U diagramu SCC se očekává, že poskytuje lepší výkonnost než diagram CUSUM na rezidua, když změna v úrovni v čase  $t = d$  produkuje velká rezidua (např. větší než 3 v absolutní hodnotě). Avšak v případě nedetekční schopnosti SCC v  $t = d$ , může být na tomto místě doporučeno použít diagram CUSUM. Takže oba tyto diagramy mají svá pro a proti při monitorování autokorelovaných procesů.

Navrhovaný regulační diagram  $\lambda_{LS,t}$  kombinuje požadované vlastnosti obou diagramů CUSUM i SCC. Dokáže zachytit jak náhlé, tak i malé změny v procesní úrovni. K ilustraci této vlastnosti regulačního diagramu  $\lambda_{LS,\max}$ , byly vybrány dva možné případy výskytu LS pro proces AR(1) s  $\phi = 0,9$  a  $\sigma = 1$ . První případ je ukázán na obrázku 5.12. V tomto případě, změna úrovně neprodukovala velká rezidua. Prvních 200 pozorování pochází z procesu AR(1) s  $\phi = 0,9$  a nulovou střední hodnotou. K posunu střední hodnoty procesu dojde v  $t = 201$ . K detekci změny procesní střední hodnoty byly zkonstruovány tři regulační diagramy – SCC, CUSUM na rezidua a  $\lambda_{LS,\max}$ . Parametry těchto tří regulačních diagramů byly vybrány tak, aby hodnota ARL ve statisticky zvládnutém stavu, byla

přibližně 370. Diagram SCC (obrázek 5.12B) nedokázal detekovat změnu v intervalu  $201 \leq t \leq 400$ . Regulační diagramy CUSUM a  $\lambda_{LS,max}$  detekovaly posun téměř ve stejný čas.

Druhý případ ukazuje realizaci diagramu  $\lambda_{LS,max}$ , když LS poskytuje velká rezidua. Časové řady na obrázku 5.13A pocházejí ze stejného procesu, jako obrázek 5.12A. Na rozdíl od předchozího případu, je vidět, že LS, která nastala v čase  $t = 201$  vytváří velká rezidua. Tato velká rezidua způsobila signalizace SCC v situaci „mimo regulační meze“. Na tomto místě, diagram CUSUM na rezidua v rychlé detekci přítomnosti LS selhal. Jak je patrné z obrázku 5.13.D, diagram  $\lambda_{LS,max}$  byl schopen zjistit posun v  $t = 201$ . Použitím případů, znázorněných na obrázcích 5.12 a 5.13 lze očekávat, že diagram  $\lambda_{LS,max}$  dosahoval lepších výsledků, než regulační diagram CUSUM na rezidua nebo SCC při monitorování autokorelovaných procesů.

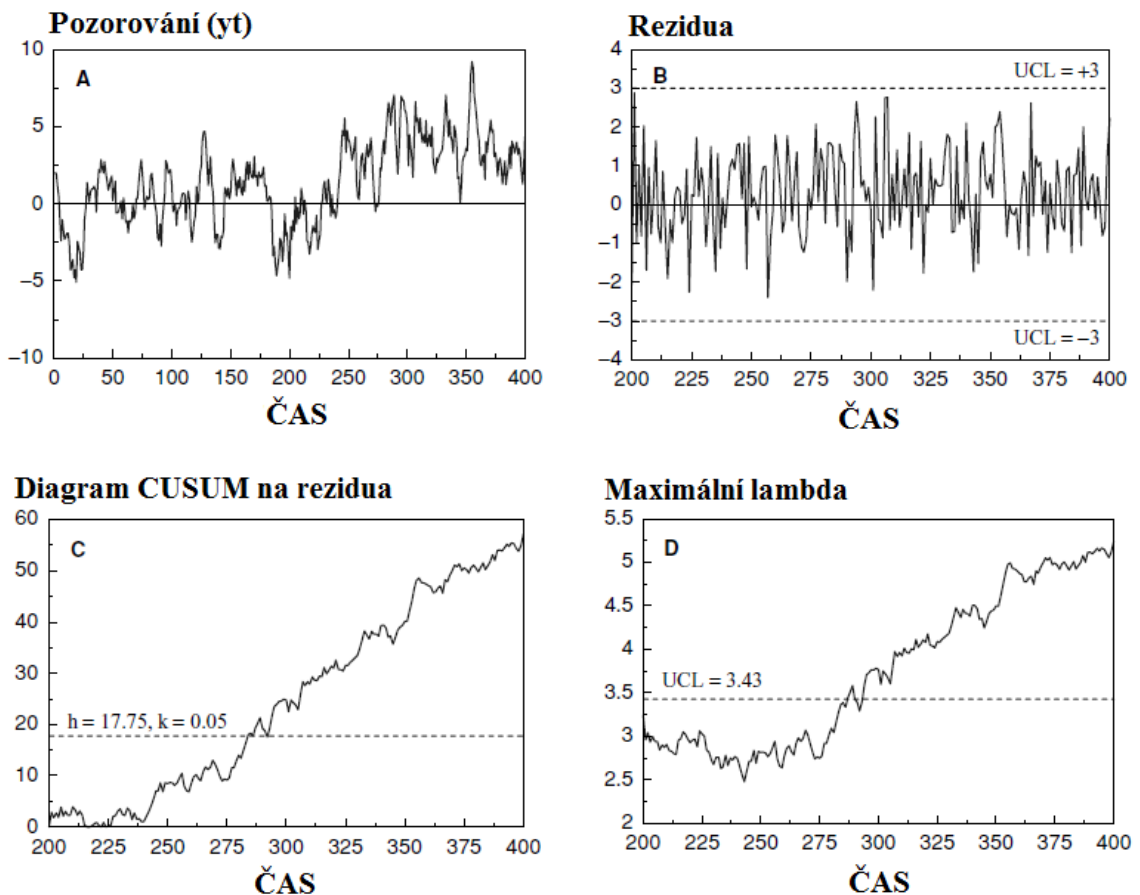
### 5.5.3 Srovnání ARL

Pro různé úrovně autokorelace  $\phi$  a posun  $\delta$  porovnáme výkonnost ARL diagramů  $\lambda_{LS,max}$  a  $\bar{\lambda}_{LS}$  s diagramem SCC, pomocí simulace Monte Carlo. Na tomto místě bych poznamenal, že posun ve střední hodnotě  $\delta$  se měří z hlediska směrodatné odchylky měření. Každý běh simulace se skládá z 5000 iterací. Simulační algoritmus byl napsán pomocí jazyka R. Proces AR(1) byl nasimulován pomocí jazyka R (zdrojový kód viz výše) a pro sekvenci i.i.d. náhodných veličin pocházejících z normálního rozdělení byl použit následující výpis v jazyku R:

```
> # simulate Gaussian White Noise process
> set.seed(123)
> y = rnorm(250)
> ts.plot(y, main="Gaussian White Noise
Process", xlab="time", ylab="y(t)",
+         col="blue", lwd=2)
> abline(h=0)
>
> # equivalent plot using plot()
> plot(y, main="Gaussian White Noise Process", type="l",
xlab="time", ylab="y(t)",
+         col="black", lwd=2)
> abline(h=0)
```

Pro účely srovnání bylo stanoveno  $m = 200$  a regulační meze diagramů  $\lambda_{LS,max}$  a  $\bar{\lambda}_{LS}$  byly stanoveny tak, aby ARL ve statisticky zvládnutém stavu odpovídalo přibližně ARL

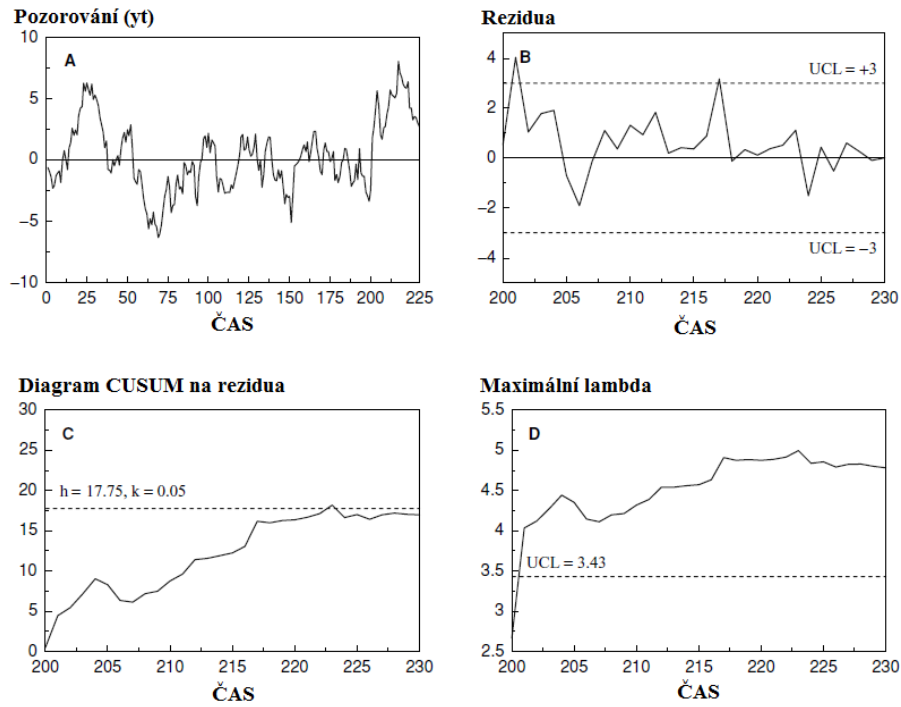
v tomto stavu, jako u diagramu SCC s mezemi  $\pm 3\sigma$ . Porovnání výsledků je uvedeno v tabulce 5.3 a zobrazeno v obrázku 5.14. Tyto grafy jasně ukazují převahu regulačního diagramu na základě  $\lambda_{LS,max}$ . Tato převaha je tím výraznější, když  $\phi > 0$ . Tyto výsledky nejsou neočekávané, protože, jak je uvedeno v předchozím odstavci, diagram  $\lambda_{LS,max}$  disponuje vhodnými vlastnostmi diagramů SCC a CUSUMu na rezidua. Je zajímavé si všimnout toho, že ve srovnání s  $\lambda_{LS,max}$ , diagram  $\bar{\lambda}_{LS}$  je výkonnější při detekci malých procesních posunů, ale zato je méně citlivý při detekci velkých posunů, zejména při velkém  $\phi$ .



**Obr. 5.12** Výkonnost diagramů SCC, CUSUM na rezidua a  $\lambda_{LS,max}$  při detekci úrovně posunu v případě, když je reziduum v bodě posunu malé

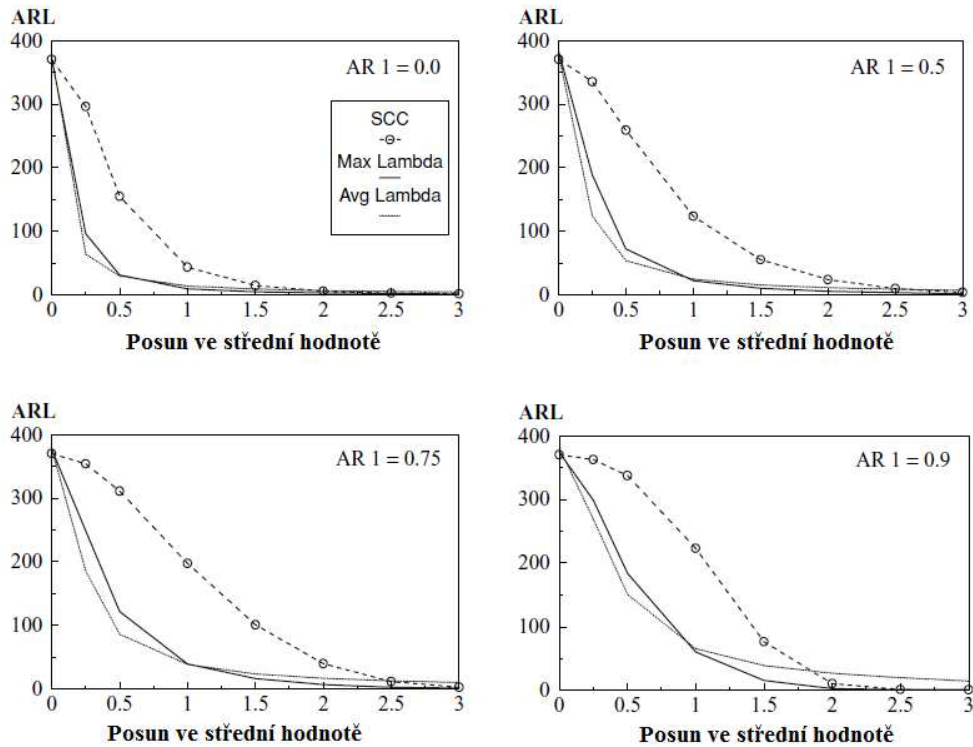
(Zdroj: Vlastní zpracování v jazyku R)





**Obr. 5.13** Výkonnost diagramů SCC, CUSUM na rezidua a  $\lambda_{LS,max}$  při detekci úrovně posunu v případě, když je reziduum v bodě posunu velké

(Zdroj: Vlastní zpracování v jazyku R)



**Obr. 5.14** Srovnání ARL: regulační diagramy SCC,  $\lambda_{LS,max}$  a  $\bar{\lambda}_{LS}$

(Zdroj: Vlastní zpracování v jazyku R)

Tab. 5.3 Srovnání výkonnosti ARL (a variability): regulační diagramy SCC,

$\lambda_{LS,max}$  a  $\bar{\lambda}_{LS}$  (Zdroj: Vlastní zpracování v jazyku R)

$\phi$	Monitorovací technika	Regulační meze	Posun ( $\delta$ )							
			0,0	0,3	0,5	1,0	1,5	2,0	2,5	3,0
0,00	SCC*	$\pm 3,00$	370,37	281,14	155,22	43,89	14,97	6,30	3,24	2,00
			(369,87)	(280,64)	(154,72)	(43,39)	(14,46)	(5,78)	(2,70)	(1,41)
	SCC**	$\pm 3,00$	373,37	296,29	162,69	43,59	14,75	6,52	0,34	2,04
			(392,44)	(287,51)	(155,78)	(41,87)	(13,97)	(5,80)	(2,72)	(1,36)
$\lambda_{LS,max}$	$\pm 3,41$	372,76	96,37	31,43	9,87	5,08	3,26	2,38	1,83	
		(380,16)	(82,06)	(21,91)	(5,83)	(2,71)	(1,60)	(1,09)	(0,78)	
$\bar{\lambda}_{LS}$	$\pm 1,64$	377,37	64,19	30,01	14,27	9,51	7,14	5,69	4,78	
		(397,93)	(46,49)	(17,57)	(7,60)	(4,74)	(3,43)	(2,67)	(2,23)	
0,25	SCC*	$\pm 3,00$	370,37	311,61	206,03	75,42	29,04	12,24	5,60	2,85
			(369,87)	(311,23)	(205,92)	(75,78)	(29,72)	(12,98)	(6,18)	(3,09)
	SCC**	$\pm 3,00$	373,37	330,40	211,33	74,24	28,95	12,39	5,74	2,88
			(392,44)	(327,35)	(198,39)	(71,76)	(29,56)	(12,70)	(6,08)	(2,94)
$\lambda_{LS,max}$	$\pm 3,44$	382,00	137,41	46,65	14,49	7,20	4,32	2,90	2,03	
		(393,14)	(124,29)	(34,43)	(9,34)	(4,31)	(2,57)	(1,65)	(1,11)	
$\bar{\lambda}_{LS}$	$\pm 1,63$	374,80	85,80	39,66	18,37	12,02	8,98	7,03	5,88	
		(396,26)	(68,30)	(25,65)	(10,24)	(6,40)	(4,60)	(3,53)	(2,86)	
0,50	SCC*	$\pm 3,00$	370,37	335,30	258,42	123,82	55,47	24,22	10,12	4,14
			(369,87)	(335,09)	(258,96)	(126,39)	(59,92)	(29,43)	(14,59)	(6,95)
	SCC**	$\pm 3,00$	373,37	346,64	270,35	115,34	53,74	22,10	9,99	4,29
			(392,44)	(362,38)	(259,48)	(117,55)	(57,58)	(26,85)	(13,47)	(7,04)
$\lambda_{LS,max}$	$\pm 3,46$	381,91	189,13	72,31	22,62	10,68	5,86	3,46	2,12	
		(397,66)	(185,13)	(60,71)	16,16)	(7,39)	(4,22)	(2,64)	(1,63)	
$\bar{\lambda}_{LS}$	$\pm 1,61$	378,32	123,88	54,01	24,81	15,87	11,53	9,00	7,42	
		(398,68)	(111,13)	(37,46)	(14,82)	(8,91)	(6,42)	(4,96)	(4,04)	
0,75	SCC*	$\pm 3,00$	370,37	354,04	311,22	197,73	101,17	40,24	11,90	3,01
			(369,87)	(354,22)	(313,63)	(210,21)	(126,48)	(68,55)	(31,01)	(11,11)
	SCC**	$\pm 3,00$	373,37	360,10	327,89	198,11	62,75	38,31	11,57	3,04
			(392,44)	(376,34)	(330,31)	(200,55)	(114,22)	(62,92)	(30,81)	(11,27)
$\lambda_{LS,max}$	$\pm 3,46$	376,78	248,75	122,06	39,56	16,53	7,25	3,06	1,46	
		(387,93)	(254,92)	(113,85)	(32,90)	(15,07)	(7,83)	(3,97)	(1,62)	
$\bar{\lambda}_{LS}$	$\pm 1,50$	378,08	185,39	86,39	38,91	23,88	17,29	13,14	10,50	
		(398,63)	(184,40)	(70,59)	(25,95)	(15,14)	(10,69)	(8,12)	(6,51)	
0,90	SCC*	$\pm 3,00$	370,37	362,63	337,59	223,30	76,44	10,68	1,40	1,00
			(369,87)	(363,96)	(347,04)	(283,60)	(169,64)	(56,90)	(10,09)	(0,97)
	SCC**	$\pm 3,00$	373,37	368,81	342,69	228,97	82,71	9,55	1,49	1,00
			(392,44)	(389,68)	(373,67)	(295,99)	(176,87)	(46,05)	(10,65)	(0,00)
$\lambda_{LS,max}$	$\pm 3,43$	374,23	298,85	183,82	60,70	15,54	2,63	1,08	1,00	
		(407,55)	(334,26)	(195,79)	(70,40)	(26,24)	(6,65)	(1,03)	(0,01)	
$\bar{\lambda}_{LS}$	$\pm 1,14$	381,08	268,98	150,95	65,65	38,99	26,84	20,05	14,85	
		(405,41)	(278,72)	(146,98)	(52,24)	(28,76)	(19,92)	(14,87)	(11,57)	

\* Vypočítáno pomocí programu od Wardell et al. (1994)

\*\* Vypočítáno pomocí simulace Monte Carlo

Většina tradičních procedur u regulačních diagramů, jsou založeny na předpokladech, že procesní měření, která jsou monitorována, jsou nezávislá a pochází ze stejného rozdělení. S příchodem vysokorychlostních systémů sběru dat, je předpoklad nezávislosti obvykle porušen. To znamená, že se autokorelace mezi měřeními stává inherentní charakteristikou stabilního procesu. Tato autokorelace způsobuje značné zhoršení výkonnosti regulačních diagramů při regulaci procesu. K vyřešení tohoto problému autokorelovaných procesů bylo navrženo několik přístupů. Nejvíce populární postup využívá buď Shewhartovův, CUSUM nebo EWMA regulační diagram pro rezidua vhodně odhadnutého ARMA modelu. Nicméně, postupy tohoto typu mají nízkou citlivost, zejména pokud se jedná o pozitivně autokorelované procesy. Jako alternativu, jsem prozkoumal používání statistik použitých v procedurách časových řad při monitorování detekce odlehlých hodnot a posunů střední hodnoty procesu.

Simulační studie se zaměřila na zjištění úrovně posunů střední hodnoty autokorelovaných procesů s důrazem na významný model AR(1). Prezentované výsledky ukázaly, že aparát pro monitorování změn v úrovni střední hodnoty autokorelovaných procesů může být založen buď na  $\lambda_{LS,max}$  nebo  $\bar{\lambda}_{LS}$ . Regulační diagram  $\lambda_{LS,max}$  disponuje vhodnými vlastnostmi jak Shewhartových diagramů, tak i CUSUM diagramu. To je právě ten důvod, proč dává vynikající výkonnost po stránce ARL ve srovnání s dosavadními postupy pro detekci změny střední hodnoty autokorelovaných procesů. Ve srovnání s diagramy SCC a  $\lambda_{LS,max}$ , je diagram  $\bar{\lambda}_{LS}$  shledáván citlivějším při detekci malých posunů, ale méně citlivý při detekci velkých procesních změn.

Velmi snadno může být navržený regulační aparát rozšířen pro účinnou identifikaci přítomnosti aditivních a inovativních odlehlých hodnot. Rafinovaná identifikace typu intervence ovlivňující proces umožní uživatelům efektivně sledovat zdroj statistické nevládnutosti procesu, která je důležitým krokem při odstraňování zvláštní příčiny variability. Je také důležité si uvědomit, že navrhovaný postup lze použít i při práci s obecnějším modelem autoregresního integrovaného klouzavého průměru (ARIMA).

Pozorování autokorelovaného procesu plynou především ze systému automatického sběru dat. Tyto systémy sběru dat jsou obvykle řízeny softwarem, který může být pro zpracování dat upgradován SPC funkcemi. V rámci takového integrovaného systému bude užitečnost navrhované procedury optimalizována.

## 6 ANALÝZA PRO VÝBĚR VHODNÉHO STATISTICKÉHO SOFTWARE

Tato kapitola je převzata z (Kovářík, 2007). Jediným předpokladem pro zavedení statistické regulace jakosti ve výrobním podniku, který používá automatický sběr dat, je výběr vhodného statistického programu. Proto jsem zpracoval analýzu vybraných softwarů, které jsem hodnotil podle vybraných parametrů. Tuto analýzu jsem doplnil cenovým přehledem těchto vybraných produktů.

Při výběru vhodného software jsem sledoval následující parametry kvality:

1. Technické požadavky zahrnují minimální kapacitu operační paměti, velikost pevného disku, nutnost matematického koprocessoru, typ grafického zobrazení a typ tiskáren a plotterů.
2. Úroveň obsluhy se týká způsobu manipulace s daty, ovládání systémů obvykle cestou hierarchických nabídek (menu) nebo ikon, interaktivní nebo dávkové zpracování dat a manipulace s výsledky.
3. Úroveň metod a algoritmů jsem posuzoval dle numerického a statistického hlediska. Je třeba si uvědomit, že každá statistická metoda platí za jistých a často dosti striktních předpokladů, a pokud nejsou tyto předpoklady splněny, jsou výsledky buď velmi přibližné, nebo zcela zavádějící a chybné.
4. Ověřování předpokladů statistických metod je proto nezbytnou součástí kvalitních programů. Navíc je často důležité volit i alternativní postupy, které jsou užitečné při nesplnění některých předpokladů, resp. pracují adaptivně, přizpůsobují se chování dat.

V současné době začíná docházet postupně k unifikaci obsluhy programových balíčků související s jejich aplikací pod operačními systémy typu WINDOWS. Rozdíly se projevují především v rozsahu nabídky různých metod a numerické stránce použitých algoritmů. Zde je kvalitativní posun velmi malý a řada nových verzí se značně zlepšenými možnostmi obsluhy je prakticky stejná s ohledem na úroveň metod a algoritmů. Ve statistických programech se vyskytují různé techniky, vedoucí často i k rozporným závěrům. Některé statistické metody jsou použitelné jen pro některé typy modelů. Zde uvádím některé programové balíky vhodné ke zpracování experimentálních dat:

- *Unistat* - Nabízí, jako jediný svého druhu, možnost přepínání jazyka uživatelského prostředí a jazyka výstupů mezi češtinou, slovenštinou a angličtinou. UNISTAT může

fungovat jako nádstavba EXCELU s přidáním nabídek UNISTATu do nabídky Excelu. Samozřejmostí je export výsledků do MS office a HTML.

- **SPSS (SPSS Inc., USA)** - Je koncipován modulárně a umožňuje práci v příkazovém a interaktivním režimu, ale také v dávkovém režimu. Umožňuje zpracování rozsáhlých datových souborů, vyskytujících se např. v ekonomii a sociologii.
- **Statistica** - Jeho moduly poskytují uživatelům celou řadu datových analýz, prostředků k řízení a správě dat, grafických zobrazení dat a data miningových procedur. Využívá nejnovějších technologií a postupů pro prediktivní modelování, shlukování, klasifikování a průzkumné techniky v jediné integrované řadě modulů.
- **Minitab (Quality Companion)** - Součástí je soubor tutoriálů a pomocníků KeepingTab. ActivStats for MINITAB je interaktivní multimediální statistický text, ve kterém jsou animace, dynamické grafy a videoklipy.
- **SAS (SAS Institute Inc., USA)** - Je orientován na komplexní užití v technické praxi, a proto má i speciální moduly, např. pro řízení kvality, pro analýzu spolehlivosti, atd.
- **Statgraphics (STSC Inc., USA)** - Vyznačuje se jednoduchostí ovládání systémem nabídek a panelů. Umožňuje kreslit a zobrazovat i velmi náročné grafy.
- **SOLO (BMDP Statistical Software Inc., USA)** - Zajímavý statistický systém v jazyce BASIC. Je řízený systémem nabídek a panelů bez předdefinovaných hodnot. V základním modulu obsahuje programy pro nejběžnější statistické úlohy a v pokročilém modulu i komplikovanější úlohy. Patří mezi nejlepší při porovnání jeho možností s jeho cenou.
- **NCSS (BMDP / SPSS, USA)** - Obsahuje výkonný tabulkový procesor podobný Excelu s kapacitou 16384 řádků a 256 sloupců, v němž lze snadno upravovat a transformovat data. Dobře je řešen import a export dat a export grafiky.
- **SYSTAT (Systat / SPSS, USA)** - Jedná se o spolehlivý software s poměrně širokým a vyváženým výběrem statistických metod, z nichž pro některé je však nutné znát některé syntaktické příkazy. Poněkud obtížnější je práce s daty a jejich úprava.
- **S-PLUS (MathSoft, USA)** - Systém nabízí zhruba 1400 statistických, matematických a dalších funkcí prakticky ze všech oblastí statistické analýzy, které jsou přístupné pomocí redefinovatelného menu, tlačítek, nebo z příkazového řádku. Přes značnou rozsáhlost systému je jeho dynamickou strukturou zajištěna neobyčejně snadná obsluha.

- **ADSTAT (TriloByte Statistical Software, Česká republika)** - Nabízí 8 základních modulů obsahujících přes 30 statistických metod. Grafické výstupy lze tisknout v prezentační kvalitě, nebo ukládat do souborů ve formátu TIF pro snadný export do jiných programových balíčků.
- **QC - EXPERT (TriloByte Statistical Software, Česká republika)** - Český statistický software pro sledování jakosti a analýzu dat v technologii, laboratoři a obchodu. Program poskytuje praktický přehledný výstup v podobě protokolu a grafů.

Následovat bude tabulka porovnání cen vybraných statistických programů, která má sloužit jako přehled dostupných statistických programů, které v sobě mají moduly týkající se SPC. Podle mého názoru neexistuje žádný univerzální statistický software, ale jako vrcholový nástroj pro statistickou analýzu jednoznačně shledávám jazyk R a Statgraphics.

**Tab. 6.1** Porovnání cen vybraných statistických programů  
(Zdroj: vlastní zpracování)

Druh Softwaru	Cena (ze 17. 5. 2013)
Unistat	23 630,- Kč <sup>6</sup>
IBM SPSS (SPSS Inc., USA)	40 000,- Kč – 139 940,- Kč <sup>7</sup>
Statistica (SPC Standard Plus)	62 835,- Kč – 93 980,- Kč <sup>8</sup>
Minitab 16	37 526,- Kč
SAS (SAS Institute Inc., USA)	895 - 3500 USD <sup>8</sup>
Statgraphics (STSC Inc., USA)	55 600,- Kč
SOLO (BMDP Statistical Software Inc., USA)	2 695 USD
NCSS (BMDP / SPSS, USA)	6 999,- Kč
SYSTAT (Systat / SPSS, USA)	21 985,- Kč <sup>8</sup>
S-PLUS (MathSoft, USA)	51 080,- Kč
ADSTAT (TriloByte Statistical Software, Česká republika)	25 000,- Kč <sup>7</sup>
QC - EXPERT (TriloByte Statistical Software, Česká republika)	45 900 – 79 700,- Kč <sup>8</sup>

<sup>6</sup> Každá další licence (dodatečná) 11 900,- Kč.

<sup>7</sup> Záleží na druhu licence, která je poskytována pro určitý počet uživatelů.

<sup>8</sup> Záleží na počtu modulů, které si klient v této verzi objedná.

## ZÁVĚR

V úvodní kapitole praktické části této práce jsem nejprve představil systém automatizovaného sběru dat v podniku Continental Barum, s.r.o., kde nyní působím jako statistik. Výstupem takových systémů je v mnoha případech časová řada autokorelovaných pozorování, což má za následek špatnou implementaci klasických Shewhartových regulačních diagramů. Následkem této skutečnosti jsou zkreslené hodnoty indexů způsobilosti a výkonnosti procesu, jako jsou například  $C_p$ ,  $C_{pk}$ ,  $P_{pk}$  a  $C_{pm}$ .

Další nosnou kapitolou praktické části (kapitola 4) tvořily ukázky monitorování autokorelovaných procesů na praktických případových studiích z reálných procesů ze své osobní praxe v oblasti SPC.

Hlavním cílem této diplomové práce bylo provést srovnání jednak aparátu, navrženého Atienzou a Tangem (2002) a aparátu navrženého Montgomerym a Friedmanem (1989) s klasickým Shewhartovým regulačním diagramem po stránce výkonnosti v simultánním monitorování střední hodnoty, rozptylu a autokorelační struktury sériově korelovaných procesů. Následně jsem se zaměřil na studium chování regulačního diagramu  $\lambda_{LS,t}$  navrženého Atienzou et al. (1998) používaného pro detekci úrovně posunu v procesu AR(1). Výkonnost tohoto aparátu byla vyhodnocena a srovnána s Alvanovým a Robertsovým aparátem SCC (diagram zvláštních příčin – modifikovaný Shewhartův regulační diagram pro rezidua). Pro srovnání výkonnosti těchto regulačních diagramů byla použita simulace Monte Carlo implementovaná v programovém prostředí jazyka R, která se nachází v knihovně `spc`.

Jedním ze způsobů, jak monitorovat autokorelaci procesu je používat tři regulační diagramy (jeden pro detekci změny střední hodnoty, druhý pro rozptyl, a třetí pro autokorelační strukturu dané řady). Jinou alternativou je vytvořit monitorovací proceduru, která může současně detekovat změny v MVAS. Pomocí vlastností stacionárního Gaussovského ARMA procesu, byl vyvinut regulační aparát založený na statistice  $H$ . Ve srovnání s ostatními Shewhartovými aparáty pro monitorování autokorelovaných procesů, je navrhovaný klouzavý diagram  $H$  účinný v simultánní detekci posunů v MVAS dané řady. Výsledky jsou prezentovány v podkapitole 5.1

Simulační studie v podkapitole 5.2 se zaměřila na zjištění úrovně posunů střední hodnoty autokorelovaných procesů s důrazem na významný model AR(1). Prezentované výsledky ukázaly, že aparát pro monitorování změn v úrovni střední hodnoty autokorelovaných

procesů může být založen buď na  $\lambda_{LS,max}$  nebo  $\bar{\lambda}_{LS}$ . Regulační diagram  $\lambda_{LS,max}$  disponuje vhodnými vlastnostmi jak Shewhartových diagramů, tak i CUSUM diagramu. To je právě ten důvod, proč dává vynikající výkonnost po stránce ARL ve srovnání s dosavadními postupy pro detekci změny střední hodnoty autokorelovaných procesů. Ve srovnání s diagramy SCC a  $\lambda_{LS,max}$ , je diagram  $\bar{\lambda}_{LS}$  shledáván citlivějším při detekci malých posunů, ale méně citlivý při detekci velkých procesních změn.

Velmi snadno může být navržený regulační aparát rozšířen pro účinnou identifikaci přítomnosti aditivních a inovativních odlehlých hodnot. Rafinovaná identifikace typu intervence ovlivňující proces umožní uživatelům efektivně sledovat zdroj statistické nevládnutosti procesu, která je důležitým krokem při odstraňování zvláštní příčiny variability. Je také důležité si uvědomit, že navrhovaný postup lze použít i při práci s obecnějším modelem autoregresního integrovaného klouzavého průměru (ARIMA).

Pozorování autokorelovaného procesu plynou především ze systému automatického sběru dat. Tyto systémy sběru dat jsou obvykle řízeny softwarem, který může být pro zpracování dat upgradován SPC funkcemi. V rámci takového integrovaného systému bude užitečnost navrhované procedury optimalizována.



## CONCLUSION

In the opening chapter of the Practical part of the thesis, I introduced an automated data-gathering system in Continental Barum, Ltd. where I currently work as a statistician. Its output is usually an autocorrelated time series which results in challenging implementation of traditional Shewhart's regulation diagrams. Values of Process fitness and performance indices such as  $C_p$ ,  $C_{pk}$ ,  $P_{pk}$ , and  $C_{pm}$  are negatively impacted as well.

Next important chapter (4) consisted of autocorrelated processes' monitoring demonstrations using real-world case studies from my own practice in the field of SPC.

Main objective of the master thesis was performance comparison of frameworks devised by Atienz and Tange (2002) and by Montgomery and Friedman (1989) with traditional Shewhart's regulation diagrams when simultaneously monitoring mean, variance, and autocorrelated structures in serially-correlated processes. Subsequently, a study of the  $\lambda_{LS,t}$  regulation diagram behaviour devised by Atienza et al. (1998) used for mean shift detection in the AR(1) process was conducted. Its performance has been evaluated and compared to Alvan's and Robert's SCC framework (special causes diagram – modified Shewhart's regulation diagram for residuals). Monte Carlo simulation in the  $\mathbb{R}$  programming language's `spc` library was utilized for measuring performance of the diagrams.

One way to monitor processes' autocorrelation is to exploit regulation diagrams for mean shift detection, variance, and autocorrelated structure of the time series. Another is to create a monitoring procedure which can detect changes in MVAS. Properties of the stationary Gaussian ARMA process served as a basis for a regulation diagram using  $H$  statistic. In relation with another Shewharts' methods for autocorrelated processes' monitoring, the  $H$  diagram is effective for simultaneous shift detections in the time series' MVAS.

A simulation study in subchapter 5.2 aimed at mean shift detection level for autocorrelated processes with emphasis on the important AR(1) model. Results showed the method for monitoring mean shift detection may be based on either the  $\lambda_{LS,max}$  or  $\bar{\lambda}_{LS}$  statistic. The former has favourable properties in both Shewhart's and CUSUM diagrams, the reason why its performance ARL-wise compared to current methods for monitoring autocorrelated processes' mean shift detections. In relation to SCC diagrams and  $\lambda_{LS,max}$ ,

$\bar{\lambda}_{LS}$  is agreed to be more sensitive for smaller shifts detections. Conversely, it's less sensitive for large shifts.

The devised framework can be easily expanded for effective identification of additive and innovative outliers. Refined identification of an intervention type influencing a particular process allows users to monitor sources of process' statistical uncontrollableness, an important step for eliminating special causes of variance. It is also important to note the procedure is further usable for a more general autoregressive integrated moving average (ARIMA) model.

Observations for autocorrelated process are sourced primarily from an automatic data-gathering system, usually governed by software upgradeable with SPC functions for data processing. In such integrated system, effectiveness of the devised procedure can be optimized.

## SEZNAM POUŽITÉ LITERATURY

- [1] ALWAN, L. C., 1992. *Effects of autocorrelation on control chart performance*. Communications on Statistics: Theory and Methods. vol. 21, 1025–1049.
- [2] ALWAN, L. C. a H. V. ROBERTS. *Time-series modelling for statistical process control*. Journal of Business and Economic Statistics. 1988, vol. 16, 87–95.
- [3] APLEY, D. W. a J. SHI, 1999. *The GLRT for statistical process control of autocorrelated processes*. IIE Transactions. vol. 31, 1123–1134.
- [4] APLEY, D. W. a F. TSUNG. *The autoregressive  $T^2$  chart for monitoring univariate autocorrelated processes*. Journal of Quality Technology. 2002, vol. 34, 80–96.
- [5] ARLT, Josef, ARLTOVÁ, Markéta a Eva, RUBLÍKOVÁ, ©2002. *Analýza ekonomických časových řad s příklady*. Vyd. 1. skripta VŠE Praha, Fakulta informatiky a statistiky. 146 s. Dostupný z:  
<http://nb.vse.cz/~arltova/vyuka/crsbir02.pdf>
- [6] ATIENZA, O. O., TANG, L. C. a B. W. ANG. *A SPC procedure for detecting level shifts of autocorrelated processes*. Journal of Quality Technology. 1998, vol. 30, 340–351.
- [7] ATIENZA, O. O., TANG, L. C. a B. W. ANG. *A CUSUM scheme for autocorrelated observations*. Journal of Quality Technology. 2002, vol. 34, 187–199.
- [8] ATIENZA, O. O., TANG, L. C. a B. W. ANG, 1997. *ARL properties of a sample autocorrelation chart*. Computers and Industrial Engineering. vol. 33, 733–736.
- [9] ATIENZA, O. O., TANG, L. C. a B. W. ANG, 2002. *Simultaneous monitoring of sample and group autocorrelations*. Quality Engineering. vol. 14, 489–499.
- [10] BALKIN, S. D. a D. K. LIN. *Performance of sensitizing rules on Shewhart control charts with autocorrelated data*. International Journal of Reliability, Quality and Safety Engineering. 2001, vol. 8, 159–171.
- [11] BOX, G. E. P., JENKINS, G. M. a G. C. REINSEL, 1994. *Time Series Analysis, Forecasting and Control*. Englewood Cliffs, NJ: Prentice Hall.
- [12] BROCKWELL, P. J. a R. A. DAVIS, 1987. *Time Series Analysis: Theory and Methods*. New York: Springer-Verlag.
- [13] CASTAGLIOLA, P. a F. TSUNG, 2005. *Autocorrelated SPC for non-normal situations*. Quality and Reliability Engineering International. vol. 21, 131–161.
- [14] ČSN ISO 7870 *Regulační diagramy: obecné pokyny a úvod*. Praha: ČNI, 1995.

- [15] DOOLEY, K. J. a S. G. KAPOOR. *An enhanced quality evaluation system for continuous manufacturing processes, Part 1: Theory*. Transactions of the ASME, Journal of Engineering for Industry. 1990, vol. 112, 57–62.
- [16] DYER, J. N., CONERLY, D. M. a B. M. ADAMS. *A simulation study and evaluation of multivariate forecast based control charts applied to ARMA processes*. Journal of Statistical Computation and Simulation. 2003, vol. 73, 709–724.
- [17] FABIAN, František, HORÁLEK, Vratislav, KŘEPELA, Josef, MICHÁLEK, Jiří, CHMELÍK, Václav, CHODOUNSKÝ Jiří a Jan KRÁL, 2007. *Statistické metody řízení jakosti*. Praha: ČSJ. ISBN 978-80-02-01897-1.
- [18] FUCHS, C. a Ron S. KENETT, 1998. *Multivariate Quality Control*. Vyd. 1. New York: Marcel Dekker, Inc. 212 s. ISBN 0-8247-9939-9.
- [19] HARRIS, T. J. a W. H. ROSS. *Statistical process control procedures for correlated observations*. Canadian Journal of Chemical Engineering. 1991, vol. 69, 48-57.
- [20] CHEN, C. a L. M. LIU. *Joint estimation of model parameters and outlier effects in time series*. Journal of the American Statistical Association. 1993, vol. 88, 284–297.
- [21] KALGONDA, A. A. a S. R. KULKARNI. *Multivariate quality control chart for autocorrelated processes*. Journal of Applied Statistics. 2004, vol. 31, 317–327.
- [22] KOVÁŘÍK, Martin, 2007. Projekt zavedení statistické regulace jakosti v podniku Tegü Vuko, s.r.o. Univerzita Tomáše Bati ve Zlíně, Fakulta managementu a ekonomiky. Diplomová práce.
- [23] KRIEGER, C. A., CHAMP, C. W. a L. C. ALWAN, 1992. *Monitoring an autocorrelated process. Presented at the Pittsburgh Conference on Modeling and Simulation*.
- [24] KUPKA, Karel, ©2001. *Metody statistického řízení jakosti*. AUTOMA. č. 7–8. s. 13–17, [online]. [cit. 2010-9-25]. Dostupný z:  
<http://automa.hjc.cz/download/au070113.pdf>
- [25] LU, C. W. a M. R. REYNOLDS. *Control charts for monitoring the mean and variance of autocorrelated processes*. Journal of Quality Technology. 1999, vol. 31, no. 3, 259–274.
- [26] LU, C. W. a M. R. REYNOLDS. *EWMA control charts for monitoring the mean of autocorrelated processes*. Journal of Quality Technology. 1999, vol. 31, 166–188
- [27] MALIARIS, A. G. *Stochastic methods in economics and finance*. Elsevier Science. 1982.

- [28] MASTRANGELO, C. M. a E. C. BROWN. *Shift detection properties of moving centerline control chart schemes*. Journal of Quality Technology. 2000, vol. 32, 67–74.
- [29] MELOUN, Milan a Jiří MILITKÝ, 2004. *Statistická analýza experimentálních dat*. Vyd. 2. Praha: Academia, nakladatelství Akademie věd České republiky. 953 s. ISBN 80-200-1254-0.
- [30] MONTGOMERY, D. C. a D. J. FRIEDMAN, 1989. *Statistical process control in computerintegrated manufacturing environment*. In J. B. Keats and N. F. Hubele (eds). *Statistical Process Control in Automated Manufacturing*. New York: Marcel Dekker.
- [31] MONTGOMERY, D. C., 2009. *Introduction to Statistical Quality Control*. Vyd. 6. USA: John Wiley & Sons, Inc. 734 s. ISBN 978-0-470-16992-6.
- [32] MONTGOMERY, D. C. a C. M. MASTRANGELO. *Some statistical process control methods for autocorrelated data (with discussion)*. Journal of Quality Technology. 1991, vol. 23, 179–204.
- [33] NOOROSSANA, R. a S. J. M. VAGHEFI, 2005. *Effect of autocorrelation on performance of the MCUSUM control chart*. Quality and Reliability Engineering International. vol. 22, 191–197.
- [34] NOSKIEVIČOVÁ, Darja, ©2008. *Vybrané metody statistické regulace procesu pro autokorelovaná data*. AUTOMA. č. 10. s. 40–43, [online]. [cit. 2010-9-17]. Dostupný z:  
[http://www.odbornecasopisy.cz/index.php?id\\_document=37926](http://www.odbornecasopisy.cz/index.php?id_document=37926)
- [35] ROBERTS, S. W., 1959. *Control chart tests Based on Geometric Moving Averages*. Technometrics. vol. 1, 239–250.
- [36] RUNGER, G. C. a T. R. WILLEMAIN, 1996. *Batch-means control charts for autocorrelated data*. IIE Transactions. vol. 28, 483–487.
- [37] SUN, J. a L. XU, 2004. *Batch average control chart*. ASQ Annual Quality Congress Proceedings. 58.
- [38] TESTIK, M. C., 2005. *Model inadequacy and residuals control charts for autocorrelated processes*. Quality and Reliability Engineering International. vol. 21, 115–130.
- [39] TIMMER, D. H., PIGNATIELLO, J. a M. LONGNECKER, 1998. *The development and evaluation of CUSUM-based control charts for an AR(1) process*. IIE Transactions. vol. 30, 525–534.
- [40] TOŠENOVSKÝ, Josef a Darja NOSKIEVIČOVÁ, 2000. *Statistické metody pro zlepšování jakosti*. Vyd. 1. Ostrava: Montanex, a.s. 362 s. ISBN 80-7225-040-X.
- [41] TSAY, R. S., 1988. *Outliers, level shifts, and variance changes in time series*. Journal of Forecasting. vol. 7, 1–20.

- [42] WARDELL, D. G., MOSKOWITZ, H. a R. D. PLANTE, 1994. *Run-length distributions of specialcause control charts for correlated processes (with discussion)*. Technometrics. vol. 36, 3–27.
- [43] WRIGHT, P. A., 1995. *A process capability index sensitive to skewness*. Journal of Statistica Computation and Simulation. vol. 52, 195 – 203.
- [44] YOURSTONE, S. A. a D. C. MONTGOMERY, 1989. *A time-series approach to discrete real-time process quality control*. Quality and Reliability Engineering International. vol. 5, 309–317.
- [45] ZHANG, N. F., 2000. *Statistical control charts for monitoring the mean of a stationary process*. Journal of Statistical Computation and Simulation. vol. 66, 249–258.

## SEZNAM OBRÁZKŮ

Obr. 1.1 Shewhartův diagram .....	16
Obr. 1.2 Výběr racionální podskupiny.....	16
Obr. 1.3 Riziko zbytečného signálu $\alpha$ a riziko chybějícího signálu $\beta$ .....	17
Obr. 1.4 Křivky operační charakteristiky pro regulační diagram $\bar{x}$ s mezemi v rozpětí $3\sigma$ . $\beta = P$ (pravděpodobnost, že nebude zjištěn posun střední hodnoty o $k\sigma$ včas) .....	21
Obr. 1.5 Hodnoty ARL pro diagram $\bar{x}$ s mezemi o rozpětí $3\sigma$ v případech posunu procesu o $k\sigma$ .....	23
Obr. 1.6 Závislost v datech způsobuje vzrůst zdánlivé variability a nepoužitelnost Shewhartova diagramu .....	25
Obr. 1.7 Autokorelace v datech může mít přiřaditelnou příčinu, po jejímž odstranění se výrazně zmenší variabilita .....	26
Obr. 1.8 Shewhartův diagram x-individual vykazuje nesmyslných 53 porušení pravidel v důsledku přítomnosti autokorelace.....	27
Obr. 1.9 Diagram EWMA, popř. diagram AR, reálně diagnostikuje pět pravidel porušení .....	27
Obr. 2.1 Srovnání diagramů z hlediska vlastností jejich paměti zobrazených pomocí průběhu váhových koeficientů $w$ (%) .....	30
Obr. 2.2 Vlastnosti paměti diagram EWMA zobrazené pomocí průběhu váhových koeficientů $w$ (%) .....	33
Obr. 3.1 Ukázka sběru dat v podniku Continental Barum, s.r.o. ....	47
Obr. 3.2 Schématické znázornění vstupů a výstupů systému automatizovaného sběru dat v podniku Continental Barum, s.r.o.....	47
Obr. 3.3 Možný výstup z QDA v podobě Shewhartova regulačního diagramu .....	49
Obr. 3.4 Automatický výpočet indexů způsobilosti procesu .....	49
Obr. 3.5 Ukázka denního reportu, který nabízí jednoduchý přehled kvality po jednotlivých dnech .....	50
Obr. 3.6 Systém QTREE-SPC C/S pro sběr a vyhodnocování procesních dat .....	51
Obr. 3.7 Možné grafické výstupy ze systému QTREE-SPC C/S .....	51
Obr. 4.1 Exploratorní analýza ukazuje, že data pocházejí z normálního rozdělení.....	55
Obr. 4.2 Regulační diagramy $\bar{x}_j$ , $R_{kl}$ pro naměřené hodnoty viskozity.....	55
Obr. 4.3 Autokorelační funkce (vlevo), autokorelace reziduí (vpravo).....	56

Obr. 4.4 Exploratorní analýza ukazuje na normalitu chyb predikce $e_k$ pro $\lambda_{opt}$ .....	57
Obr. 4.5 Dynamický diagram EWMA pro naměřené hodnoty viskozity .....	59
Obr. 4.6 Diagram “x individual” .....	59
Obr. 4.7 Kvantilový graf potvrzuje normalitu dat (vlevo). Data s vyznačeným trendem a vyhlazením (vpravo).....	60
Obr. 4.8 Graf autokorelace s mezemi významnosti korelačních koeficientů .....	60
Obr. 4.9 Dynamický diagram EWMA s jednokrokovou predikcí.....	61
Obr. 4.10 Znázornění časové řady měření šířky patního pásku u bočnice .....	62
Obr. 4.11 ACF a PACF původní časové řady.....	63
Obr. 4.12 Periodogram původní časové řady .....	63
Obr. 4.13 Graf reziduí a reziduální periodogram modelu ARIMA(0,1,0).....	63
Obr. 4.14 Reziduální ACF a PACF modelu ARIMA(0,1,0) <sub>c</sub> .....	64
Obr. 4.15 Reziduální ACF a PACF modelu ARIMA(1,1,0) <sub>c</sub> .....	64
Obr. 4.16 Reziduální ACF a PACF modelu ARIMA(0,1,1) <sub>c</sub> .....	65
Obr. 4.17 Reziduální ACF a PACF modelu ARIMA(1,1,1) <sub>c</sub> .....	66
Obr. 4.18 Časová řada s předpověďmi .....	67
Obr. 4.19 Regulační diagram pro individuální hodnoty aplikovaný na rezidua zvoleného ARIMA modelu (vlevo) a regulační diagram EWMA pro rezidua identifikovaného modelu ARIMA s parametrem $\lambda = 0,2$ (vpravo).....	68
Obr. 4.20 Regulační diagram CUSUM.....	68
Obr. 4.21 Křivky ARL pro změny střední hodnoty u diagramů EWMA (vlevo) a CUSUM (vpravo) .....	68
Obr. 5.1 Vliv změny n na citlivost MH diagramu při detekci posunu střední hodnoty.....	74
Obr. 5.2 Chování ARL diagramů SCC, SACC a MH při detekci posunu procesní střední hodnoty .....	75
Obr. 5.3 Citlivost diagramů SCC, SACC a MH při detekci posunu ve variabilitě.....	76
Obr. 5.4 Citlivost grafů SCC, SACC a MH při detekci posunů pro různé parametry modelu .....	76
Obr. 5.5 Simulovaná měření daného procesu .....	77
Obr. 5.6 Diagram MH pro řadu z předchozího obrázku .....	78
Obr. 5.7 (A) Nasimulovaná řada AR(1) s $\phi = 0,5$ . (B) Nasimulovaná řada AR(1) s AO v $t = 51$ . (C) Nasimulovaná řada AR(1) s IO v $t = 51$ . (D) Nasimulovaná řada AR(1) s LS v $t = 51$ .....	83



Obr. 5.8 Simulovaná časová řada AR(1) ( $\phi = 0,9$ , $\sigma^2 = 1,0$ a $\omega_{LS,101} = \sigma_z$ .....	89
Obr. 5.9 Diagram zvláštních příčin pro řady z obrázku 23.2.....	89
Obr. 5.10 Vlastnosti a chování $\lambda_{LS,t}$ pro $t = 1, \dots, T = 100$ (bez posunu).....	91
Obr. 5.11 Vlastnosti a chování $\lambda_{LS,t}$ pro $t = 1, \dots, T = 100$ (úroveň posunu velikosti $\omega_{LS,101} = \sigma_z$ v $t = 101$ ) .....	91
Obr. 5.12 Výkonnost diagramů SCC, CUSUM na rezidua a $\lambda_{LS,max}$ při detekci úrovně posunu v případě, když je reziduum v bodě posunu malé.....	96
Obr. 5.13 Výkonnost diagramů SCC, CUSUM na rezidua a $\lambda_{LS,max}$ při detekci úrovně posunu v případě, když je reziduum v bodě posunu velké.....	97
Obr. 5.14 Srovnání ARL: regulační diagramy SCC, $\lambda_{LS,max}$ , a $\bar{\lambda}_{LS}$ .....	97

**SEZNAM TABULEK**

Tab. 4.1 Vstupní údaje vypočtené hodnoty EWMA <sub>k-1</sub> a e <sub>k</sub> pro $\lambda = 0,56$ .....	54
Tab. 4.2 Hodnoty sumy čtverců chyb predikce $\sum_{k=1}^{50} e_k^2$ pro různé hodnoty $\lambda$ .....	56
Tab. 4.3 Hodnoty střední čáry a regulačních mezí dynamického diagramu EWMA pro $\lambda_{opt.} = 0,56$ .....	58
Tab. 4.4 Šířka patního pásku u bočnice v $1 \cdot 10^{-1}$ mm.....	62
Tab. 4.5 Výstupní tabulka pro model ARIMA(0,1,0) <sub>c</sub> .....	64
Tab. 4.6 Výstupní tabulka modelu ARIMA(1,1,0) <sub>c</sub> .....	65
Tab. 4.7 Výstupní tabulka modelu ARIMA(0,1,1) <sub>c</sub> .....	66
Tab. 4.8 Výstupní tabulka pro model ARIMA(1,1,1) <sub>c</sub> .....	67
Tab. 5.1 Simulovaný proces měření, $196 \leq t \leq 205$ .....	78
Tab. 5.2 Výpočet charakteristiky $\lambda_{LS,t}$ , pomocí MS Excel 2010.....	90
Tab. 5.3 Srovnání výkonnosti ARL (a $\sigma$ ): regulační diagramy SCC, $\lambda_{LS,max}$ , a $\bar{\lambda}_{LS}$ .....	98
Tab. 6.1 Porovnání cen vybraných statistických programů.....	102

## SEZNAM PŘÍLOH

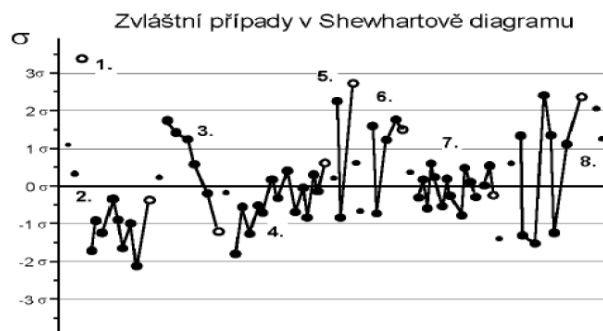
- Příloha P I: Typická porušení předpokladů pro aplikaci regulace Shewhartovými diagramy v různých technologiích
- Příloha P II: Nejčastěji používané testy nenáhodných seskupení
- Příloha P III: Členění regulačních diagramů dle vybraných hledisek

**PŘÍLOHA P I: TYPICKÁ PORUŠENÍ PŘEDPOKLADŮ PRO  
APLIKACI REGULACE SHEWHARTOVÝMI  
DIAGRAMY V RŮZNÝCH TECHNOLOGIÍCH**

<b>Předpoklady pro aplikaci regulace Shewhartovými diagramy v různých technologiích</b>				
<b>Odvětví, technologie (veličina)</b>	<b>Normalita</b>	<b>Nezávislost</b>	<b>Konstantní střední hodnota</b>	<b>Homogenita</b>
Strojírenství (rozměry)	ano	ano	ano	ano
Mechanické zkušebny (pevnost, pružnost)	ne	ano	ano	ne
Chemie, metalurgie, hutnictví (koncentrace, obsahy)	ano	ne	ne	ne
Chemie, metalurgie, hutnictví (ostatní fyzikální parametry)	ano	ne	ne	ne
Životní prostředí (koncentrace)	ne	ne	ne	ne
Elektrické veličiny, součástky	ano	ano	ano	ne
Energetika	ne	ne	ne	ne
Plasty, polymery, textil (fyzikálně mechanické veličiny)	ne	ano	ne	ano
Biochemie, farmacie, potravinářství	ne	ne	ano	ano
Vnitropodnikové ekonomické a finanční ukazatele	ne	ne	ne	ano
Sociologie, lidské zdroje	ne	ne	ne	ne

*(Zdroj: Meloun a Militký, 2004)*

## PŘÍLOHA P II: NEJČASTĚJI POUŽÍVANÉ TESTY NENÁHODNÝCH SESKUPENÍ



- Pravidlo 1.** *Jedna hodnota je mimo regulační meze.*  
Lokální porucha procesu, chybné měření, výpadek. Chybně stanovené regulační meze, malá variabilita uvnitř podskupiny při konstrukci diagramu. Opakuje-li se na téže straně, může jít o posunutí střední hodnoty nebo asymetrické rozložení dat. Opakuje-li se na obou stranách může jít o zvýšení nestability nebo rozptylu dat.
- 
- Pravidlo 2.** *9 hodnot je na téže straně od centrální linie.*  
Pravděpodobné posunutí střední hodnoty, snížení variability mezi podskupinami, asymetrie dat, příliš široké nebo neodpovídající regulační meze.
- 
- Pravidlo 3.** *6 hodnot monotónně roste či klesá.*  
Autokorelovaný proces, závislá měření. Lineární trend, způsobuje opotřebením nebo výpadkem. Příliš široké regulační meze. Odstraněním přiřaditelné příčiny lze někdy zvýšit  $C_p$ .
- 
- Pravidlo 4.** *14 alternujících hodnot.*  
Přeregulovaný nebo nestabilní proces. Autokorelovaná měření se záporným  $r$ . Odstraněním přiřaditelné příčiny lze někdy zvýšit  $C_p$ . Podvádění operátorem, vymyšlená čísla.
- 
- Pravidlo 5.** *2 ze 3 hodnot mimo interval  $\pm 2s$ .*  
Varování před možným překročením regulačních mezí.
- 
- Pravidlo 6.** *4 z 5 hodnot mimo intervall  $\pm s$  na téže straně centrální linie.*  
Pravděpodobné posunutí střední hodnoty. Varování před možným překročením regulačních mezí.
- 
- Pravidlo 7.** *15 hodnot je uvnitř intervalu  $\pm s$ .*  
Snížení variability mezi podskupinami. Při opakování uvažovat o nových regulačních mezích. Nesprávná volba regulačních mezí. Podvádění operátorem, vymyšlená čísla.
- 
- Pravidlo 8.** *8 hodnot je mimo interval  $\pm s$  na obou stranách centrální linie.*  
Zvýšení variability mezi podskupinami. Varování před překročením regulačních mezí. Porucha procesu.

### PŘÍLOHA P III: ČLENĚNÍ REGULAČNÍCH DIAGRAMŮ DLE VYBRANÝCH HLEDISEK

<i>Hledisko</i>	<i>Typy regulačních diagramů</i>
Počet regulačních mezí	<ul style="list-style-type: none"> <li>- regulační diagramy pro jednostrannou regulaci</li> <li>- regulační diagramy pro oboustrannou regulaci</li> </ul>
Charakter regulované veličiny	<ul style="list-style-type: none"> <li>- regulační diagramy pro SPC měření</li> <li>- regulační diagramy pro SPC srovnáváním</li> </ul>
Počet znaků jakosti simultánně sledovaných na 1 jednotce ve výběru	<ul style="list-style-type: none"> <li>- regulační diagramy pro sledování jednoho znaku jakosti (klasické Shewhartovy diagramy)</li> <li>- regulační diagramy pro sledování více znaků jakosti najednou (např. Hotellingův diagram)</li> </ul>
Stupeň opakovatelnosti procesu	<ul style="list-style-type: none"> <li>- regulační diagramy pro proces s vysokým stupněm opakovatelnosti (klasické Shewhartovy diagramy)</li> <li>- regulační diagramy pro procesy s nízkým stupněm opakovatelnosti (např. cílové a standardizované diagramy)</li> </ul>
Zohlednění předchozích hodnot výběrové charakteristiky ve výpočtu aktuální hodnoty výběrové charakteristiky	<ul style="list-style-type: none"> <li>- regulační diagramy bez paměti (klasické Shewhartovy diagramy)</li> <li>- regulační diagramy s pamětí (např. diagramy CUSUM, EWMA)</li> </ul>
Použité testové kritérium	<ul style="list-style-type: none"> <li>- regulační diagramy pro monitorování polohy procesu (např. diagram <math>\bar{x}</math>)</li> <li>- regulační diagramy pro sledování stejnoměrnosti procesu (např. diagram <math>R, s</math>)</li> <li>- regulační diagramy pro sledování počtu, resp. podílu neshodných jednotek (např. diagram <math>p</math>)</li> <li>- regulační diagramy pro sledování počtu, resp. podílu neshod (např. diagram <math>c</math>)</li> </ul>
Zohlednění rizika zbytečného signálu $\alpha$ a rizika chybějícího signálu $\beta$	<ul style="list-style-type: none"> <li>- regulační diagramy pracující pouze s rizikem <math>\alpha</math> (klasické Shewhartovy diagramy)</li> <li>- regulační diagramy zohledňující oba druhy rizika</li> </ul>
Závislost hodnot regulované veličiny	<ul style="list-style-type: none"> <li>- regulační diagramy pro nezávislá data (např. klasické Shewhartovy diagramy)</li> <li>- regulační diagramy pro závislá data (např. diagram reziduí, dynamický EWMA diagram)</li> </ul>