

# Metody práce s autokorelovanými daty v SPC bez použití modelu (Model – free approach to autocorrelated data in SPC)

Jan Fridrich

Tento příspěvek se zabývá tématem srovnání metod statistické regulace procesu pro autokorelovaná data. Text je postaven na porovnání dvou základních přístupů k práci s autokorelovanými daty, kterými jsou přístup matematického modelování a přístup tzv. bez modelu. Hlavní část práce je věnována přístupu bez modelu a zejména metodice průměrů z podskupin, která je v současnosti nejrozvinutější metodikou bez použití matematického modelu chování procesu. Tato část práce je koncipována jako popis hledání alternativy k zavedeným postupům modelování, které jsou časově i zdrojově velice náročné. Závěr patří samotnému porovnání obou metod práce s autokorelovanými daty v rámci SPC. Kritéria porovnání souvisí s použitím metod v reálných podmínkách výroby, a proto byla vybrána kritéria: snadnost aplikace, účinnost v odstranění vlivu autokorelace v datech a míra poskytované informace o chování sledovaného procesu.

## ÚVOD

Navzdory obecně se periodicky opakující vlně zvyšujícího se a zase opadnuvšího zájmu o statistické řízení kvality (SQM), v každé filozofii řízení a zlepšování kvality je pro něj místo jako pro silný nástroj, jak rozhodovat o budoucím směru postupu při zlepšování kvality výrobků i procesů. Jednou z nejnápadnějších metod SQM je statistická regulace procesu (SPC), která se uplatňuje na poli statistického řízení a zlepšování procesů, kam dále patří například navrhování experimentů nebo tzv. inženýrství kvality [3]. Cílem metody je pomocí sledování chování procesu odstranit ty příčiny variability, které činí výstupy z procesu nepředvídatelnými. Tato nepředvídatelnost se projevuje skrze parametry polohy a rozptylu, jejichž hodnoty se v čase mění a činí tak proces statisticky nestabilní. Charakteristiky takového procesu nejsou robustní, nevypovídají dostatečně o úrovni dosahované kvality a nemohou být základem pro efektivní nastavení cílů zlepšování. První funkcí SPC je tedy napomoci nastolení a udržení statisticky stabilního procesu. Druhou funkcí SPC je pak funkce informační, kdy podoba vynesných statistik v regulačním diagramu ukazuje operátorům průběh výsledků kontroly kvality. Ti potom mají šanci vhodným zásahem včas zabránit výskytu neshodných produktů a jejich průniku dále po proudu výroby.

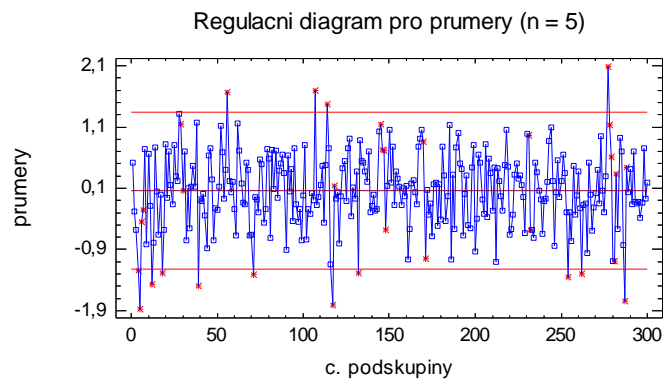
Při návrhu regulačního diagramu, který je nástrojem pro rozlišení obou druhů příčin variability, je tedy nutné navrhnout takovou podobu, která efektivně oddělí oba druhy příčin variability při současné minimalizaci rizik chybějícího nebo zbytečného signálu nestability. Takto bude zaručen stav, kdy nebudou sestavovány rozsáhlé projekty při řešení běžných operativních problémů, a na druhou stranu nenastane situace, kdy soustavným seřizováním a operativními zásahy bude léčen chybný návrh systému výroby ( tj. běžné podmínky procesu). Takový zbytečný zásah znamená likvidaci shodných produktů, zbytečné zásahy údržby, vede k porušení vyhovujícího stavu procesu a ve výsledku znamená ztrátu času, materiálu i dalších zdrojů a další náklady na zásah a odstranění nekvalitních produktů.

## 1 VLIV AUTOKORELACE NA NÁVRH REGULAČNÍHO DIAGRAMU

Pro efektivní návrh regulačního diagramu hraje důležitou roli otázka, zda ve výrobním systému nepůsobí vlivy způsobující vzájemnou závislost mezi získanými daty (tj. zda nejsou výstupní data z procesu autokorelována). Fenomén autokorelace dat se především týká těch výrob, kde se provádí měření v procesech automaticky, dochází k častému vzorkování resp. výběrům z procesu a dále v podmínkách určitých specifických procesů. Typickými představiteli procesů s autokorelovanými výstupními daty jsou procesy velkosériové a hromadné výroby s použitím automatické kontroly a procesy spojitě výroby v chemickém, metalurgickém a potravinářském průmyslu. V těchto odvětvích se často objevují tzv. vnitřní faktory výroby, jako

jsou reaktory, recyklační toky nebo čerpání materiálu ze zásobníků (tanků), které mohou svým charakterem způsobovat závislost ve výstupních datech.

Jistá míra autokorelace je pro data obecně vrozená a také některé regulační diagramy jako CuSum a EWMA této vlastnosti využívají. Vysoká autokorelace však způsobuje zkreslení průběhu dat vůči regulačním mezím diagramu a vede tak k chybným rozhodnutím ohledně zásahů do procesu. Projevem působení silné autokorelace je obvykle vysoký výskyt bodů mimo regulační meze jak ukazuje následující příklad (viz obr. 1), ve kterém je zobrazen regulační diagram pro výběrové průměry.



Obr. 1 Regulační diagram pro průměry dvojice  $\bar{x} - s$

Z regulačního diagramu na obr. 1 je patrné, že se v průběhu dat vůči regulačním mezím vyskytuje nápadně mnoho signálů působení vymezených příčin variability. V této fázi je jasné, že prezentovaný regulační diagram není schopen efektivně oddělit vymezené a náhodné příčiny variability.

## 2 ODSTRANĚNÍ VLIVU AUTOKORELACE NA NÁVRH REGULAČNÍHO DIAGRAMU

Pro odstranění vlivu autokorelace v datech lze použít dva základní přístupy:

- a) metody s použitím modelu (Model – based approach),
- b) metody bez modelu (Model – free approach).

Při aplikaci metody s použitím modelu se hledá matematický model, který vysvětlí chování procesu, tj. také vlastnosti působení vlivů autokorelace. Mezi tyto matematické modely se nejčastěji počítá modelování časových řad a jeho aproximace pomocí EWMA statistiky [4]. Po nalezení modelu, který nejlépe vystihuje chování procesu, se hledají reziduální odchylky od tohoto modelu podle vztahu  $\varepsilon_t = x_t - \hat{x}_t$ , kde je  $x_t$  empirická hodnota a  $\hat{x}_t$  je hodnota odhadnutá modelem. O těchto odchylkách od modelu chování se předpokládá, že mají následující vlastnosti:

1.  $E(\varepsilon_t) = 0$ ,
  2.  $D(\varepsilon_t) = \sigma^2$ ,
  3.  $\text{cov}(\varepsilon_t; \varepsilon_{t-1}) = 0$ ,
  4.  $\varepsilon_t \sim N(0; \sigma^2)$ .
- (1)

Reziduální odchylky tak splňují všechny předpoklady o datech pro efektivní použití regulačních diagramů, a proto na ně lze aplikovat tradiční Shewhartovy regulační diagramy.

Metody s použitím modelu jsou nejpřesnějším odhadem chování procesu, ale jejich použití je náročné jak na čas, tak na úroveň znalostí o statistických metodách. Především nutnost znalosti pokročilých statistických metod pak nutí podniky k dodatečným výdajům na získání

kvalifikovaného pracovníka, popř. na další výcvik současných pracovníků v použití statistických metod.

Pro tyto prezentované nevýhody metod založených na matematickém modelu se řada zahraničních autorů rozhodla hledat alternativu, která by v praxi nahradila jejich použití. V minulosti vznikla řada takových metod bez modelu, ale jen následující tři byly dále rozpracovány:

- a) metoda kontextuálních regulačních diagramů [9],
- b) metoda DFTC (Distribution Free Tabular CuSum) [7],
- c) metoda průměru z podskupin (Batch means – BM) [1].

Z těchto tří zde bude prezentována právě metoda BM, která je v současnosti nejrozvinutější metodou bez použití modelu.

### 3 METODA PRŮMĚRŮ Z PODSKUPIN

Metoda průměrů z podskupin (Batch means – BM) je postavena na čtyřech parametrech:  $n$ ,  $b$ ,  $w$  a  $k$ . Parametr  $n$  představuje obecně velký výběrový soubor, který tvoří jedno dlouhé pozorování procesu v řádu tisíce hodnot.

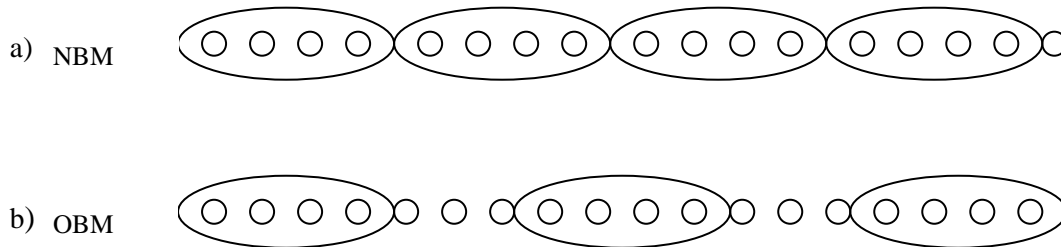
Výběrový soubor  $n$  je dále rozdělen na  $k$  podskupin velikosti  $b$  s tím, že mezi jednotlivými podskupinami je vynecháno obecně  $w$  jednotek, které nebudou započítány do analýzy. Z vytvořených podskupin jsou pak vypočteny průměry, které se dále vynášejí do regulačního diagramu. Změnou oproti klasickému regulačnímu diagramu pro průměry je nakládání s BM jako s individuálními hodnotami (použije se Shewhartův regulační diagram pro individuální hodnoty). Právě nakládání se získanými průměry jako s individuálními hodnotami zaručuje odstranění vlivu autokorelace při analýze regulačního diagramu. Jelikož vliv autokorelace je obecně vnímán jako součást náhodných příčin variability, způsobuje i tento vliv především změny v krátkodobé variabilitě uvnitř podskupiny [4]. Pokud se tedy jedná s vypočtenými průměry jako s individuálními hodnotami, omezuje se vliv této variability a klade se důraz na dlouhodobou variabilitu mezi podskupinami.

V předchozí kapitole bylo naznačeno, že postupy bez použití modelu by měly být alternativou ke složitým postupům použití matematických modelů. Zde se však naráží na omezení metody BM, kterým je mimo předpokladu prostředí bohatého na data také předpoklad stacionární časové řady původních dat. Přestože v žádné z dostupné literatury [1], [4] o BM není přímo řečeno, že metoda BM není vhodná pro nestacionární časové řady, ve všech případech bylo jejich použití prezentováno na simulacích dat obecně z ARMA ( $p; q$ ) modelu. Jedinými autory, kteří připouští možnost použití BM pro všechny případy časových řad, jsou Smeti, Koronakis a Kousourdis [8], jež však problém ponechávají jen v rovině obecného předpokladu o BM. Přestože se tedy v žádném textu explicitně nevyklučuje možnost použití metody BM pro nestacionární modely časových řad, lze spíše usuzovat, že se její použití omezuje jen na rodinu ARMA ( $p; q$ ) modelů.

Vstupem do analýzy statistické stability procesu v podmínkách autokorelovaných dat je tedy soubor dat velikosti  $n$ , jehož jednotky vynesené v průběhovém diagramu jsou časovou řadou stacionární v čase. Výběrový soubor  $n$  je dále rozdělen do podskupin podle následujícího schématu:

1. Vytvoří se první podskupina obvyklým způsobem z prvních  $b$  hodnot,
2. Další podskupiny se pak vytvářejí (obr. 2):
  - a) z následujících hodnot a tedy bez vynechání resp. přeskočení ( $w = 0$ ) (tzv. Non-overleapping batch means NBM),

b) přeskočením počtu  $w$  jednotek a následným vytvořením další podskupiny velikosti  $b$  (Overleapping batch means OBM).



Obr. 2 Zobrazení způsobu tvorby podskupin dle NBM a OBM

Pro rozdělení souboru do podskupin je třeba určit počet podskupin  $k$ , počet vypuštěných jednotek  $w$  a velikost podskupiny  $b$ . Správné nadefinování těchto parametrů zajistí nejen, že průměry z podskupin budou vzájemně nezávislé, ale také asymptoticky normálně rozdělené, tj. bude možné aplikovat princip centrální limitní věty. Základem úspěšného použití BM metody je ale především určení velikosti podskupiny  $b$ . Jejím určením se zabývala řada autorů. Například Runger a Willemain [1] tvrdí, že pro odstranění autokorelace v datech je nejlepší velikost podskupiny asi 60 jednotek metodou NBM, zatímco Alexopoulos, Seila a Fishman [5] navrhuje velikost podskupiny maximálně 10 až 30 jednotek s tím, že se použije pro rozdělení výběrového souboru metoda OBM. Posledními autory jsou Alwan a Radson [2], kteří navrhuje velikost  $b = 4 - 6$  jednotek také metodou OBM. Autoři Alexopoulos, Seila a Fishman [5] navíc nabízí jednu z metod přímého odhadu velikosti  $b$ . Velikost  $b$  lze získat následujícími dvěma způsoby[6]:

$$\text{a) } b = n/k \quad (2)$$

$$\text{b) } b = k = \sqrt{n} \quad (3)$$

První způsob počítá s tím, že je známa velikost výběrového souboru a je rozhodnuto o počtu potřebných podskupin. Tento postup obecně vede k nadhodnocení variability odhadu polohy procesu. Druhý způsob vychází jen ze znalosti velikosti výběrového souboru, ale v tomto případě je vliv variability odhadu procesu podhodnocen a tak stejně jako v předchozím způsobu výpočtu je odhad polohy procesu nekonzistentní. V praxi lze využít pro určení velikosti  $b$  oba způsoby, ale v další práci s vypočtenými průměry je třeba počítat s omezeními obou metod.

Jiný způsob určení  $b$  navrhuje autoři Runger a Willemain [3], kdy na počátku je  $b = 1$  a v každé další iteraci se velikost  $b$  zdvojnásobí. Tento postup pokračuje do chvíle, než autokorelační koeficient v časové vzdálenosti  $h = 1$  klesne na požadovanou hodnotu  $\rho_1 \leq 0,1$ . Tento způsob je pravděpodobně pro praxi nejjednodušší a ve výsledku nejefektivnější, protože se očekává nejmenší velikost  $b$ , která je už schopna snížit vliv autokorelace [1].

Poslední zde prezentovaný způsob navrhuje Song [6], která vytvořila vztah pro určení  $b$  tak, aby byla redukována nejen autokorelace prvního řádu, ale také všech dalších řádů (resp. časových vzdáleností  $h$ ). Vztah pro určení  $b$  má tuto podobu:

$$b = 1 + [2n(c_b^2 / c_v)(\gamma_1 / \gamma_0)^2]^{1/3}, \quad (4)$$

kde je:

$c_b$ a $c_v$ ...	konstanty závislé na způsobu tvorby bodového odhadu polohy, pro OBM je to (1 a 4/3) a pro NBM jsou rovny (1 a 2),
$\gamma_0$ ...	je počet autokorelovaných jednotek potřebných pro kvalitní odhad srovnatelný s výsledky z nekorelovaných dat, který je vztažen na jednotku těchto nekorelovaných dat. (tj. jaký počet jednotek korelovaných dat je potřeba pro stejně konzistentní odhad na jednotku nekorelovaných dat),
$\gamma_1$ ...	suma všech vážených korelačních koeficientů.

Proměnné  $\gamma_0$  a  $\gamma_1$  se získají ze vztahů:

$$\gamma_0 = \lim_{n \rightarrow \infty} (1 + 2 \sum_{h=1}^n \rho_h) = 1 + 2 \sum_{h=1}^{\infty} \rho_h, \quad (5)$$

$$\gamma_1 = \lim_{n \rightarrow \infty} (2 \sum_{h=1}^n h \rho_h) = 2 \sum_{h=1}^{\infty} h \rho_h, \quad (6)$$

kde je:

$h$ ...	vzdálenost dvou hodnot v čase,
$\rho_h$ ...	autokorelační koeficient v časové vzdálenosti $h$ .

Metoda dovoluje odstranit autokorelaci v celém spektru časových vzdáleností, ale je třeba počítat i s určitým omezením. Velkou nevýhodou tohoto postupu je vysoká citlivost na průběh autokorelační funkce původních dat, kdy metoda poskytuje jinou hodnotu  $b$  pro každé opakování analýzy. Výsledkem je potřeba navrhovat jedinečný regulační diagram pro každé opakování a tedy potřeba automatizace celého postupu.

S velikostí podskupiny  $b$  dále souvisí také počet jednotek vynechaných z analýzy  $w$ . Pokud se rozhodneme pro metodu OBM, lze  $w$  najít opět iterativním způsobem, kdy se kombinují různé velikosti  $b$  s různými velikostmi  $w$  nebo se zafixuje jedna velikost  $b$  například na hodnotě 10 až 30 jednotek [5] a  $w$  se hledá iterativním způsobem, kde  $w_{i+1} = 2 \cdot w_i$ .

#### 4 ANALÝZA VÝKONNOSTI BM APLIKACÍ NA REÁLNÁ VÝROBNÍ DATA

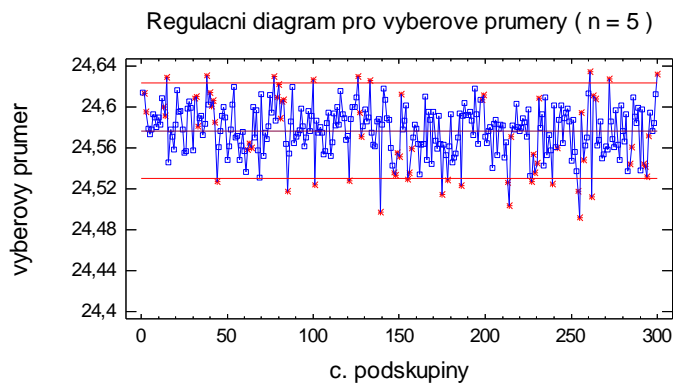
Výkonnost jednotlivých metod určení velikosti  $b$  bude prezentována na souboru dat z procesu reálné výroby v potravinářském průmyslu.

Analýza výkonnosti jednotlivých metod kopíruje předpokládaný algoritmus, jak se bude pracovat s metodou BM v reálných podmínkách průmyslových procesů. Celý postup analýzy probíhá v těchto několika krocích:

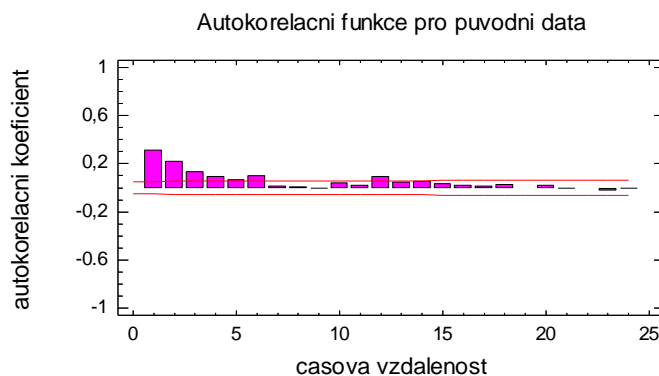
1. Prezentace vlivu autokorelace v datech aplikací Shewhartova regulačního diagramu pro výběrové průměry a charakteru působení autokorelace pomocí grafu autokorelační funkce.
2. Prezentace výsledků snížení autokorelace pro různé metody BM a různé velikosti  $b$  a  $w$  v jedné souhrnné tabulce. Postup hledání  $b$  a  $w$  bude podobný jako je předpokládaná aplikace metody BM v podmínkách reálného provozu, tj. od nejjednodušší metody po nejsložitější. Tento algoritmus probíhá v následujících krocích:
  - a) Nejjednodušším způsobem určení  $b$  je metoda NBM s  $b = k = \sqrt{n}$ , kdy v tomto případě je  $b = \sqrt{1500} = 39$ .

- b) Pokud by tato velikost  $b$  nebyla schopna zajistit dostatečné snížení autokorelace, pak použijeme metodu NBM s iterativním postupem hledání  $b$ , kdy  $b_{i+1} = 2 \cdot b_i$ .
  - c) Je – li způsob NBM neúčinný při odstranění závislosti, pak lze přejít ke způsobu OBM, kdy v tomto případě je  $b$  zafixováno na  $b = 20$  a iterativně se hledá optimální  $w$ , které bude postupně nabývat hodnot 2, 4, 8 a 16.
  - d) Poslední metodou je určení  $b$  dle Song, která je nejsložitější pro výpočet, ale počítá s autokorelací ve všech vzdálenostech  $h \geq 1$ .
3. Na základě dosažených výsledků snížení  $\rho_I$ , dosažení všeobecné nezávislosti dat a dodržení normality rozdělení je vybrána vhodná metoda BM a velikost  $b$  pro snižování závislosti v datech. Nezávislost získaných průměrů z podskupin a jejich normalita rozdělení jsou testovány pomocí testů hypotéz.
  4. Prezentace grafu autokorelační funkce pro průměry z podskupin vzniklé vybranou metodou BM a také regulační diagram pro individuální hodnoty, navržený na základě těchto průměrů z podskupin.
  5. Komentář dosaženého snížení autokorelace a vlivu na informaci o chování procesu, které lze získat z analýzy průběhu statistik v regulačním diagramu.

Vybraný soubor dat pochází z procesu hromadné výroby a sestává z 1490 hodnot měření klíčového parametru výrobku. Pro tato data byl sestaven Shewhartův regulační diagram pro výběrové průměry (obr. 3) a byly provedeny testy nezávislosti.



Obr. 3 Regulační diagram pro výběrové průměry původních dat



Obr. 4 Graf autokorelační funkce pro původní data

Tab. 1 – výsledky ověření nezávislosti

Typ testu	P - hodnota
Test založený na počtu jednotek nad a pod mediánem	0,000
Test založený na počtu zvrátů	0,037
Box - Pierceův test	0,000

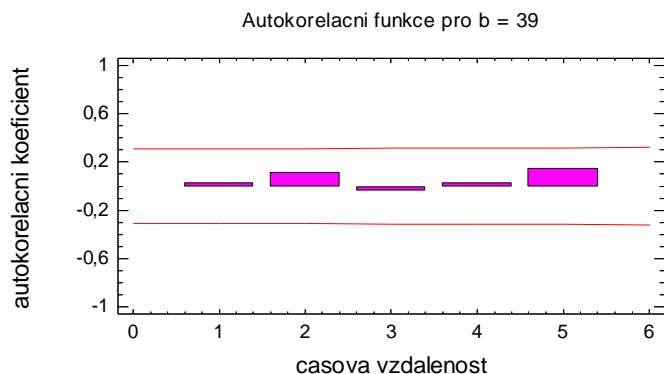
Přestože autokorelace  $\rho_1$  v datech není vysoká (viz obr.4), způsobuje řadu falešných signálů v průběhu dat v regulačním diagramu (obr. 3). Následující tabulka (tab.2) zachycuje, jak se jednotlivé metody vypořádaly s jejím působením.

Tab. 2 Výsledky snižování autokorelace pro různé metody BM

	metoda hledání b	$\rho_1$	ověření normality (p - hodnoty)				ověření normality (p - hodnoty)		
			X <sup>2</sup> - test	Shapiro - Wilkův	test šikmosti	test špičatosti	Počet jednotek nad a pod mediánem	Znaménkový test	Box - Pierceův test
a)	<b>b=39</b>	<b>0,021</b>	<b>0,000</b>	<b>0,024</b>	<b>0,989</b>	<b>0,786</b>	<b>0,827</b>	<b>0,003</b>	<b>0,995</b>
b)	<b>b=2</b>	0,325	0,100	0,714	0,066	0,001	0,000	0,002	0,000
	<b>b=4</b>	0,306	0,110	0,188	0,240	0,908	0,000	0,000	0,001
	<b>b=8</b>	0,223	0,333	0,894	0,670	0,304	0,023	0,259	0,009
	<b>b=16</b>	0,108	0,835	0,381	0,525	0,502	0,604	0,150	0,400
	<b>b=32</b>	0,070	0,816	0,918	0,912	0,993	0,456	0,597	0,850
	<b>b=64</b>	0,201	0,154	0,309	0,633	0,304	0,531	0,557	0,955
	<b>b=128</b>	0,208	0,154	0,309	0,633	0,304	0,762	1,000	0,743
c)	<b>w=2</b>	0,095	0,000	0,009	0,204	0,410	0,579	0,013	0,665
	<b>w=4</b>	-0,014	0,000	0,019	0,665	0,346	0,129	0,724	0,817
	<b>w=8</b>	-0,050	0,000	0,004	0,400	0,241	0,529	0,702	0,727
	<b>w=16</b>	0,068	0,000	0,042	0,975	0,291	0,876	0,236	0,982
d)	<b>b=26(Song)</b>	0,152	0,000	0,037	0,854	0,448	0,000	0,000	0,000

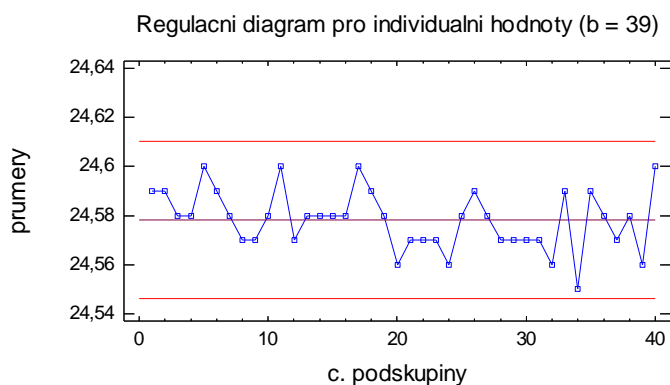
Z navržených metod snižování autokorelace v datech byla vybrána jako nejvhodnější metoda NBM s  $b = \sqrt{n} = 39$ , která účinně snižuje závislost v datech ( $\rho_1 = 0,021$ ) a navíc je jednoduchá pro aplikaci. Metoda OBM s  $b = 20$  a  $w = 4$  je ještě účinnější ve snižování autokorelace, ale dochází při jejím použití k vypouštění dat z analýzy, ztrátě informace o chování procesu a k působení negativní zbytkové autokorelace.

Výkonnost metody při snižování závislosti v datech (viz. obr. 5) a regulační diagram pro vzniklé průměry (obr. 6) z podskupin ukazují následující grafy.



Obr. 5 Autokorelační funkce pro NBM s  $b = 39$

Jak plyne z grafu autokorelační funkce (obr. 5), metoda NBM je velice účinná ve snížení autokorelace ve všech časových vzdálenostech. Průběh vzniklých nezávislých průměrů z podskupin zachycuje následující diagram (obr. 6).



Obr. 6 Regulační diagram pro BM velikosti  $b = 39$

Regulační diagram pro individuální hodnoty, který staví na průměrech podskupin velikosti 39 jednotek, ukazuje, že proces je statisticky stabilní a predikovatelný. Na data dále nepůsobí žádná forma autokorelace.

Další tabulka (tab. 3) přináší souhrnné porovnání výhod a nevýhod obou způsobů práce s autokorelovanými daty.



Tab. 3 Shrnutí charakteristik různých přístupů práce se závislými daty

Metoda	Výhody	Nevýhody
<b>Přístup s použitím modelu</b>		
<b>ARIMA</b>	<ul style="list-style-type: none"> <li>- přesný popis působení vlivů autokorelace</li> <li>- aplikovatelné na všechny druhy dat a podoby autokorelace</li> </ul>	<ul style="list-style-type: none"> <li>- časově náročné</li> <li>- náročné na zdroje</li> <li>- vyžaduje znalost pokročilých statistických metod</li> <li>- citlivost odhadu parametrů modelu na dostatek dat</li> <li>- náročná aplikace v případě sledování více proměnných</li> </ul>
<b>Aproximace modelu časové řady pomocí EWMA statistiky</b>	<ul style="list-style-type: none"> <li>- snadná aplikace</li> <li>- snadná automatizace</li> </ul>	<ul style="list-style-type: none"> <li>- citlivost odhadu parametrů modelu na dostatek dat</li> <li>- předpoklad modelu ARIMA(0;1;1)</li> <li>- požadavek doplnění o průběhový diagram jednotlivých hodnot</li> </ul>
<b>Dynamická EWMA</b>	<ul style="list-style-type: none"> <li>- bez nutnosti dalšího diagramu</li> <li>- informace o statistické stabilitě i dynamice procesu v jednom diagramu</li> </ul>	<ul style="list-style-type: none"> <li>- citlivost odhadu parametrů modelu na dostatek dat</li> <li>- předpoklad modelu ARIMA(0;1;1)</li> </ul>
<b>Přístup bez modelu</b>		
<b>Prodloužení kontrolního intervalu</b>	<ul style="list-style-type: none"> <li>- snadná aplikace</li> <li>- účinnost snížení závislosti dat</li> </ul>	<ul style="list-style-type: none"> <li>- velká ztráta informace o chování procesu</li> </ul>
<b>NBM</b>	<ul style="list-style-type: none"> <li>- snadnost implementace, výrazně se neliší od obecného postupu SPC</li> <li>- účinnost snížení autokorelace</li> <li>- příznivý poměr mezi snížením závislosti a ztrátou dat</li> <li>- snadná automatizace</li> </ul>	<ul style="list-style-type: none"> <li>- velikost podskupin při silné autokorelaci</li> <li>- platí pro vybraný časový interval</li> <li>- neúčinné pro silně autokorelovaná data ve vzdálenosti <math>h &gt; 1</math></li> </ul>
<b>OBM</b>	<ul style="list-style-type: none"> <li>- účinnost snížení autokorelace v datech</li> <li>- příznivý poměr mezi snížením závislosti a ztrátou dat</li> <li>- snadná aplikace</li> </ul>	<ul style="list-style-type: none"> <li>- ztráta informace o chování procesu</li> <li>- obraz procesu necelistvý</li> <li>- časová náročnost hledání vhodné kombinace <math>b</math> a <math>w</math></li> <li>- velikost podskupin při silné autokorelaci</li> <li>- neúčinné pro silně autokorelovaná data ve vzdálenosti <math>h &gt; 1</math></li> </ul>

## ZÁVĚR

V konceptu neustálého zlepšování kvality výrobků a procesů zastávají metody a nástroje SQM důležitou roli pro efektivní rozhodování o dalším postupu. Základním krokem zlepšování je vždy posouzení současného stavu a na jeho základě nastavení vhodných cílů zlepšování. Aby však bylo toto nastavení účinné, je třeba zajistit robustnost a predikovatelnost odhadu parametrů polohy a variability procesu.

Tento požadavek však vážně narušuje porušení předpokladu pro analýzu predikovatelnosti, kterým je předpoklad nezávislosti získaných dat. Zvýšená závislost (autokorelace) mezi daty způsobuje vážný nárůst rizika obou chyb, které působí při analýze predikovatelnosti (statistické stability) použitím regulačního diagramu.

Existují dva přístupy, jak odstranit vliv autokorelace při analýze statistické stability, a to přístup tzv. s modelem a bez modelu. První z nich staví na vypracování určitého matematického modelu chování procesu a použití reziduálních odchylek tohoto modelu jako základu pro analýzu predikovatelnosti procesu. Tento postup vyžaduje mimo regulačního diagramu pro rezidua také průběhový diagram individuálních hodnot jako záznam přirozené dynamiky procesu. Hledání modelu a tvorba obou diagramů je časově i zdrojově náročné, a proto se řada autorů rozhodla hledat alternativní postup bez použití modelu.

Mezi metody bez použití modelu je dnes nejlépe rozpracována pro praxi a pravděpodobně také nejjednodušší pro aplikaci metoda průměrů z podskupin (BM). Ve srovnání s přístupy s modelem přináší BM především hodnotu ve své jednoduchosti aplikace, kdy požadavek na rozsah souboru  $n$  je u obou přístupů srovnatelný.

Z výsledků srovnání v rámci zpracované diplomové práce vyplynulo, že metoda BM je vhodná pro ty situace, kdy není třeba tolik přesný popis chování procesu. Naopak přístup pomocí matematického modelování se uplatní tehdy, kdy jsou výstupní data z procesu silně autokorelována nebo je třeba detailní popis chování procesu.

## LITERATURA

- [1] RUNGER, George C.; WILLEMAIN, Thomas R. Batch-means control charts for autocorrelated data. *IIE Transactions*. 1995, 28, s. 483 - 487.
- [2] ALWAN, LC; RADSON, D. TIME-SERIES INVESTIGATION OF SUBSAMPLE MEAN CHARTS. *IIE TRANSACTIONS*. 1992, 24, 5, s. 15.
- [3] MILITKÝ, Jiří; MELOUN, Milan. *Kompendium statistického zpracování dat*. Praha : Academia, 2006. Kontrola a řízení jakosti, s. 982. ISBN 80-200-1396-2.
- [4] RUNGER, George C. Batch-means charts for autocorrelated data. *IIE Transactions*. 1996, 6, 28, s. 483 - 487.
- [5] MONTGOMERY, Douglas C. *Introducion to statistical quality control*. NJ : John Wiley and Sons, 2005. SPC with autocorrelated process data, s. 759. ISBN 0-471-65631-3.
- [6] ALEXOPOULOS, Christos; FISHMAN, George S. ; SEILA, Andrew F. COMPUTATIONAL EXPERIENCE WITH THE BATCH MEANS METHOD. *Proceedings of the 1997 Winter Simulation Conference*. 1997, ., s. 194-201.
- [7] SONG, Wheyming Tina . On the estimation of optimal batch sizes in the analysis of simulation output. *European Journal of Operational Research*. 1996, 88, s. 304-319.
- [8] KIM, Seong-Hee, et al Simulation analysis : A DISTRIBUTION-FREE TABULAR CUSUM CHART FOR CORRELATED DATA WITH AUTOMATED VARIANCE ESTIMATION. In *Proceedings of the 2008 Winter Simulation Conference*. USA : IEEE, 2008. s. 417 - 425.
- [9] SMETI, Eleni; KORONAKIS, Demetrios; KOUSOURIS, Leonidas. Statistical process techniques on water toxicity data, 2004
- [10] IRAD, Ben-Gal; GAIL, Morag; ARMIN, Shmilovici. Contex - based statistical process ocontrol: a monitoring procedure for state - dependent processes. *Technometrics*. 2003, 45, 4, s. 293 - 311.