
ZÁPADOČESKÁ UNIVERZITA V PLZNI

Fakulta elektrotechnická
Katedra materiálů a technologií

DIPLOMOVÁ PRÁCE

Statistická analýza dat se zaměřením na statistickou regulaci
procesu

Autor práce: **Bc. Martin Kopsa**
Vedoucí práce: **Ing. Lukáš Kupka, Ph.D.**

ZÁPADOČESKÁ UNIVERZITA V PLZNI

Fakulta elektrotechnická
Akademický rok: 2022/2023

ZADÁNÍ DIPLOMOVÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení: **Bc. Martin KOPSA**
Osobní číslo: **E20N0010P**
Studijní program: **N0713A060011 Materiály a technologie pro elektrotechniku**
Téma práce: **Statistická analýza dat se zaměřením na statistickou regulaci procesu**
Zadávající katedra: **Katedra materiálů a technologií**

Zásady pro vypracování

1. Vypracujte přehled norem, které se zabývají oblastí statistické regulace a proveďte rozdělení všech druhů regulačních diagramů vzhledem k oblasti jejich použití.
2. Popište postupy základní statistické analýzy pro vyhodnocování dat. Určete specifika analýzy v případě aplikace na statistickou regulaci.
3. Analyzujte problémy, které vznikají v případě nenormálního rozdělení analyzovaných dat a používaná řešení.
4. Vypracujte přehled statistických SW, které lze pro uvedenou analýzu dat použít a jejich případné výhody/nevýhody.
5. Proveďte základní statistickou analýzu dodaných výsledků materiálových zkoušek a vyhodnoťte výsledky ze statistické regulace. Analýzu proveďte alespoň dvěma SW.

Rozsah diplomové práce: **40 – 60**
Rozsah grafických prací: **dle doporučení vedoucího**
Forma zpracování diplomové práce: **elektronická**

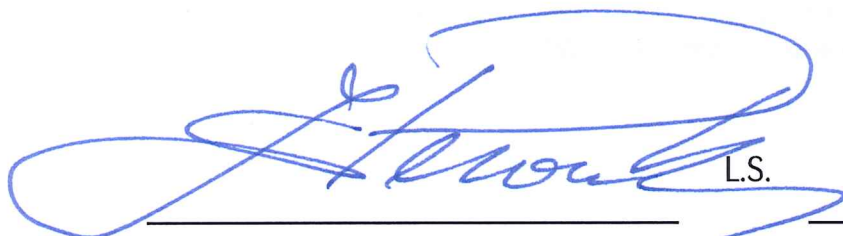
Seznam doporučené literatury:

- 1 Tůmová, O.; Pirich D.: Nástroje řízení jakosti a základy technické diagnostiky, ZČU, 2003
2. Kožíšek, J.: Statistická analýza, nakladatelství ČVUT, 1997
3. Meloun, M.; Militký, J.: Statistické zpracování experimentálních dat, East Publishing, 1998
4. manuály k SW

Vedoucí diplomové práce: **Ing. Lukáš Kupka, Ph.D.**
Katedra materiálů a technologií

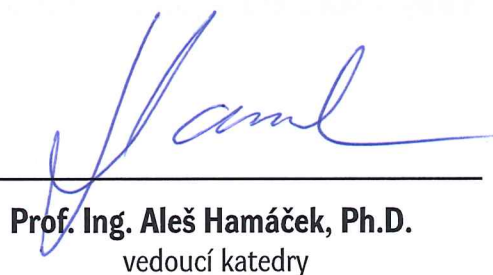
Oponent diplomové práce: **Ing. Petr Netolický, Ph.D.**
Research and Innovation Centre for Electrical Engineering

Datum zadání diplomové práce: **7. října 2022**
Termín odevzdání diplomové práce: **26. května 2023**



L.S.

Prof. Ing. Zdeněk Peroutka, Ph.D.
děkan



Prof. Ing. Aleš Hamáček, Ph.D.
vedoucí katedry

V Plzni dne 7. října 2022

Abstrakt

Předmětem této diplomové práce je statistická analýza dat se zaměřením na statistickou regulaci procesu. Práce byla rozdělena do dvou hlavních částí – na část teoretickou a část praktickou.

Teoretická část nejprve shrnuje formou rešerše základní statistické metody, které jsou důležité pro pochopení navazujících úseků práce. Dále je prostor věnován stěžejnímu tématu, statistické regulaci procesu. Kromě jeho definování vycházejícího z různých zdrojů jsou zde zanalyzovány jednotlivé etapy od přípravy až po vyhodnocení. Nedílnou součástí tohoto vyhodnocení jsou regulační diagramy. Tato část práce shrnuje jejich typy a podmínky aplikace, přičemž nejvíce pozornosti je věnováno Shewartovým regulačním diagramům. V závěru této části je popsáno nenormální rozdělení dat, statistický software a zároveň je zde uveden přehled norem souvisejících s problematikou danou zadáním.

Praktická část se zabývá vyhodnocením dat z reálného výrobního procesu s využitím znalostí z teoretické části. Vstupem jsou data, kde sledovanou hodnotou je teplota skelného přechodu. Nejprve je provedena prvotní analýza v softwaru Excel, její výsledky slouží mimo jiné i jako vstup pro další analýzu v softwaru Statistica.

Klíčová slova

Statistická regulace procesu, regulace, výrobní proces, výroba, nenormální rozdělení dat, regulační diagramy, Statistica software

Abstract

The subject of this thesis is Statistical data analysis with a focus on statistical process control. Due to the nature of the work, it was divided into a Theoretical and a Practical part.

The theoretical part in the form of research first summarizes the basic statistical methods that are important for understanding the subsequent parts of the work. The conclusion is devoted to the main topic of statistical process control. In addition to the definition of this term from various sources, attention is paid to explaining the individual stages from preparation to evaluation. Control diagrams are also an integral part of this evaluation. This work summarizes their types and application conditions, with the most attention being paid to Shewart's control charts. At the end of this section, non-normal distribution of data, statistical software and an overview of standards related to the issue given by the assignment are discussed.

The practical part is already devoted to the evaluation of data from the real production process, also with the use of knowledge from the theoretical part. The input is data, where the monitored value is the temperature of the bilge transition. First, a preliminary analysis is carried out in the Excel software that serves also as input for further analysis in the Statistica software.

Key Words

Statistical process control, control, production process, production, non-normal data distribution, control charts, Statistica software

Poděkování

Na tomto místě bych chtěl poděkovat vedoucímu práce Ing. Lukáši Kupkovi, Ph.D., a to za jeho metodické vedení práce, vstřícnost, trpělivost a věcnost.

Obsah

Úvod	- 10 -
A. Teoretická část	- 11 -
A.1 Využití metod popisné statistiky	- 11 -
A.1.1 Histogram	- 11 -
A.1.2 Statistické charakteristiky	- 11 -
A.2 Analýza dat	- 13 -
A.2.1 Korelační analýza	- 13 -
A.2.2 Statistická hypotéza	- 13 -
A.2.3 Statistická regulace procesu	- 14 -
A.3 Statistická regulace procesu	- 14 -
A.3.1 Etapy statistické regulace	- 15 -
A.3.2 Regulační diagramy	- 16 -
A.3.3 Volba typu regulačního diagramu	- 23 -
A.4 Nenormalita dat	- 26 -
A.4.1 Grafické testy normality	- 26 -
A.4.2 Početní testy normality	- 26 -
A.4.3 Práce s nenormálním rozdělením a jeho příčiny	- 27 -
A.5 Statistický software	- 31 -
A.5.1 Tíbcó statistica	- 31 -
A.5.2 IBM SPSS	- 32 -
A.5.3 Stata	- 32 -
A.5.4 SAS	- 33 -
A.5.5 Matlab	- 34 -
A.6 Normy	- 36 -
B. Praktická část	- 38 -
B.1 Práce v softwaru Excel	- 38 -
B.1.1 Společné vyhodnocení všech operátorů	- 41 -
B.1.2 Analýza hlavních operátorů	- 44 -
B.1.3 Analýza zbývajících operátorů	- 49 -

B.1.4	Vyhodnocení vlivu velikosti dávky.....	- 51 -
B.2	Analýza dat v softwaru Statistica.....	- 53 -
C.	Zhodnocení a závěr.....	- 57 -
D.	Literatura	- 59 -
E.	Přílohy	I
E.1	Vyhodnocení vlivu velikosti dávky	I

Seznam symbolů a zkratek

Metody popisné statistiky

X	soubor hodnot
\bar{x}	aritmetický průměr souboru hodnot X
x_i	i -tý prvek v souboru hodnot X
n	počet hodnot v souboru hodnot
\tilde{x}	medián
v, R	variační rozpětí
x_{max}	maximální hodnota v souboru hodnot X
x_{min}	minimální hodnota v souboru hodnot X
s_x	směrodatná odchylka souboru hodnot X
s_x^2	rozptyl souboru hodnot X
HVH	horní vnitřní hradba
DVH	dolní vnitřní hradba
\bar{x}_{75}	horní kvartil
\bar{x}_{25}	dolní kvartil

Analýza dat

r	Pearsonův korelační koeficient
cov	kovariance
p	počet stupňů volnosti
t	test významnosti korelačního koeficientu
α_v	hladina významnosti
H_0	nulová hypotéza
H_1	alternativní hypotéza

Statistická regulace procesu

α	chyba prvního druhu
β	chyba druhého druhu
μ_0	poloha centrální přímky
UCL	horní regulační mez
LCL	dolní regulační mez
WL	výstražná mez regulačního diagramu
σ	vzdálenost mezí od centrální přímky

S_i	kumulovaný součet
\bar{X}_j	pozorování v j -té řadě
λ	je váha určující citlivost na rychlou změnu střední hodnoty (parametr „míry zapomínání“)
K	konstanta pro stanovení regulačních mezí
σ_{EWMA}	je směrodatná odchylka
σ_0	požadovaná úroveň směrodatné odchylky regulované veličiny
Testy normality	
u_i	normované normální rozdělení
Fx_i	hodnota distribuční funkce teoretického normálního rozdělení
Fnx_i	hodnota empirické distribuční funkce
D_α	tabelizovaná hodnota Kolmogorov-Smirnova testu
z_i	hodnota distribuční funkce standardizovaného normálního rozdělení
y	transformovaná data
x	původní data (před transformací)
λ	parametr měřítka
η, γ	tvarové parametry
ε	parametr polohy
g	geometrický průměr

Úvod

Základní činností podniku je výroba, která znamená spojení výrobních faktorů za účelem získání výrobků ve formě produktů nebo služeb. Výroba tvoří „nadmnožinu“ výrobního procesu a dále všech činností, které jsou nezbytné pro zajištění výrobních faktorů (práce, půdy a kapitálu).

Hammer a Champy [1] výrobní proces v podniku chápou jako soubor činností, který vyžaduje jeden nebo více druhů vstupů a tvoří výstup, který má pro zákazníka hodnotu. Jak bylo uvedeno výše, výstupem můžeme chápat zhotovení výrobku nebo poskytování služeb. Cílem výrobního procesu by ale měly být pouze takové produkty nebo služby, které lze „realizovat“ na trhu a získat tak přiměřené výnosy. [2]

Paní Noskievičová ve své literatuře [3] popisuje, že na výrobní proces působí řada vnějších vlivů, které způsobí variabilitu tohoto procesu pro jinak stejné vstupy. Tato variabilita je příčinou nestejněměrnější výroby a větší pravděpodobnosti výskytu neshodných jednotek¹⁾. Příмым důsledkem tohoto jevu jsou:

- zvýšené náklady na kontrolu a zkoušení,
- větší poruchovost procesu,
- větší podíl odpadu z výroby a výrobků vyžadující přepracování.

Aby výsledkem výrobního procesu byly jen takové výrobky, které lze „realizovat“ na trhu s odpovídajícími výnosy, je důležité, aby výrobní proces byl prováděn co nejefektivněji, a to zejména s ohledem na optimální spotřebu výrobních vstupů, přiměřených nákladech a nejvhodnější volbě výrobních postupů. Jako řešení se přirozeně nabízí snížit výše zmíněnou variabilitu procesu, tedy co nejvíce eliminovat vnější vlivy působící na proces. [1], [3]–[13]

Pro zajištění těchto předpokladů je nezbytné být o výrobním procesu informován, pozorovat a analyzovat jeho průběh a případně ho regulovat. Pro tuto analýzu lze využít statistické nástroje, a to od těch nejjednodušších, až po ty složitější. K těm jednodušším patří metody popisné statistiky, naopak ke složitějším se počítá například metoda statistické regulace, jež je zároveň hlavním tématem této práce.

¹⁾ Neshodná jednotka – „zmetek“ – nevyhovující výrobek z důvodu libovolné výrobní vady. Neshoda – je to spíše kaz, který ještě nutně nemusí znamenat, že výrobek je na vyhození – např. počet bodových vad na pneumatice, v minulosti počet vadných pixelů na TV.

A. Teoretická část

A.1 Využití metod popisné statistiky

Pro dosažení určité míry spolehlivosti analyzovaných dat a s tím související následné objektivní a seriózní posouzení skutkového stavu věci, dle kterého jsou prováděny zásahy do výrobního procesu, je třeba provést jistý počet pozorování. Je namístě zmínit, že pozorování bývá označeno jako náhodný pokus, výsledek náhodného pokusu je nazýván jako náhodný jev. Konkrétně za náhodný pokus může být prohlášeno pozorování např. výskytu zmetku na výrobní lince a náhodným jevem je zjištění, že konkrétní výrobek je zmetkovitý. Výsledkem pozorování je množina hodnot neboli základní soubor dat n . Tato množina může být nekonečná. Výběrový soubor pak tvoří konečnou podmnožinu základního souboru dat. [4]–[9], [11]–[15]

A.1.1 Histogram

Histogram je grafické znázornění četností pomocí grafu sloupcového typu u souboru hodnot rozděleného do tříd. Pokud informace rozdělíme do tříd, dosáhneme jejich zhuštění, ale i určité míry ztráty. Histogram nám pak pomůže soubor hodnot posoudit například z hlediska normality dat, symetrie a výskytu odlehlých hodnot. Rozdělení relativních četností hodnot v histogramu se řídí zákonem rozdělení pravděpodobnosti. Může být normální, rovnoměrné, exponenciální. Z hlediska průběhu může být histogram plochý nebo dvouvrcholový. [4]–[9], [11]–[15]

A.1.2 Statistické charakteristiky

A.1.2.1 Charakteristiky polohy

Charakteristiky polohy vyjadřují úroveň statistického souboru z hlediska velikosti.

Aritmetický průměr

Problémem průměru je, že neudává nejčastější a ani nejpravděpodobnější hodnotu, a pokud se v souboru dat vyskytují extrémní, pak průměr značně zkresluje. [4]–[9], [11]–[15]

Aritmetický průměr je typ střední hodnoty, který lze nejjednodušeji popsat vzorcem:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (1)$$

Kvantil

Medián, kvartil a kvartil, to vše představuje kvantily. Rozdělují části, které mají přibližně stejný počet prvků. Medián představuje prostřední člen řady vzestupně uspořádaných hodnot. Jeho výpočet je odlišný, pokud soubor dat obsahuje sudý nebo lichý počet hodnot.[4]–[9], [11]–[15]

*A.1.2.2 Charakteristiky měnlivosti*Variační rozpětí

Variační rozpětí představuje rozdíl mezi největší a nejmenší hodnotou v souboru dat a udává meze, ve kterých se pohybují sledované hodnoty. Nevýhodou je možné ovlivnění odlehlými hodnotami. Pravděpodobnost výskytu odlehlých hodnot je přímo úměrná počtu hodnot v souboru dat.[4]–[9], [11]–[15]

$$v = x_{max} - x_{min} \quad (2)$$

Směrodatná odchylka a rozptyl

Obě veličiny určují, jak moc jsou hodnoty rozptýleny (odchýleny) od průměru hodnot.

$$S_x = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}} \quad (3)$$

$$S_x^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} \quad (4)$$

A.2 Analýza dat

V předchozích odstavcích byly shrnuty postupy základního vyhodnocení naměřených dat. Pro komplexnější představu o dějích probíhajících ve výrobním procesu je nutné znát jednak vztahy mezi jednotlivými sledovanými (tedy změřenými) veličinami, ale i těsnost (závislost) těchto vztahů. Těmito vztahy se rozumí vzorec, ve kterém figurují sledovaná veličina jako závislá proměnná a vnější vliv jako nezávislá proměnná (tzv. funkční nebo regresní analýza), a těsností (resp. mírou těsnosti) se rozumí korelovanost obou veličin. [4]–[9], [11]–[15]

A.2.1 Korelační analýza

Dvě proměnné jsou korelované, jestliže určité hodnoty jedné proměnné mají tendenci se vyskytovat společně s určitými hodnotami druhé proměnné. Korelace je charakterizovaná Pearsonovým korelačním koeficientem, který je v intervalu $\langle -1; 1 \rangle$ a určuje míru lineární závislosti. Pro výpočet koeficientu je nutné zajistit náhodný výběr ze souboru dat, min. jedna hodnota ze souboru dat musí mít normální rozdělení a nesmí obsahovat odlehlé hodnoty. [4]–[9], [11]–[15]

$$r = \text{cov}(x, y) / s_x \cdot s_y \quad (5)$$

Test významnosti korelačního koeficientu je nutné vždy provádět při výpočtu korelačního koeficientu. Při větším počtu dat stačí menší hodnota koeficientu pro statistickou významnost. [4]–[9], [11]–[15]

$$t = \frac{|r|\sqrt{p}}{\sqrt{1-r^2}} \quad (6)$$

$$p = n - 2 \quad (7)$$

t je porováno s tabulkovým t_{krit} a za podmínky $t > t_{krit}$ je korelační koeficient statisticky významný. Analogicky pokud $t < t_{krit}$ korelační koeficient statisticky významný není.

A.2.2 Statistická hypotéza

Statistická hypotéza spadá do oblasti statistického usuzování a používá se k ověření našich predikcí o procesu. K závěrům dodává i zhodnocení o tom, jak jsou tyto závěry spolehlivé. Staví proti sobě ověřovanou nulovou hypotézu H_0 a alternativní hypotézu H_1 . Jednostrannými (pouze pro formulaci H_1) nebo dvoustrannými (formulace H_0) statistickými testy prověřujeme platnost (resp. splnění podmínky) nulové hypotézy na vzorku náhodného výběru. Pokud výsledky splňují počáteční podmínku, nezamítáme

nulovou hypotézu. Pokud výsledky splňují alternativní hypotézu, zamítáme H_0 a přijímáme H_1 . U každého závěru je nutné uvést hladinu významnosti. Rozhodování o nulové hypotéze na hladině významnosti na základě náhodného výběru může vést ke vzniku dvou možných chyb. Pro chybu 1. druhu platí nulová hypotéza, ale my ji zamítáme. Rozhodnutí o zamítnutí je dáno hladinou významnosti. K chybě 2. druhu dochází, pokud nepoznáme, že nulová hypotéza neplatí, tzn., že nulová hypotéza sice neplatí, ale my ji nezamítáme. Pravděpodobnost, s jakou neplatnost hypotézy objevíme, se nazývá síla testu. [4]–[9], [11]–[15]

A.2.3 Statistická regulace procesu

Paní Tůmová [16] ve svých skriptech popisuje statistickou regulaci jako nejvhodnější způsob kontroly mezioperační hromadné a sériové výroby.

Principy SPC byly zformulovány již ve 20. letech minulého století W. A. Shewhartem. Podstatou SPC je bezprostřední a pravidelná výběrová kontrola kvality procesů. Statistická regulace má preventivní charakter, umožňuje zásah do výroby ještě před skutečným výskytem zmetků. Výrobní proces je tak trvale udržován v ustáleném stavu. Časový rozvrh kontroly, způsob výběru vzorků ke kontrole a rozhodování o jakosti se řídí předpisy založenými na základech matematické statistiky. Tím se tento přístup liší například od další možnosti operační kontroly namátkové, která pouze konstatuje jakost dávky po provedené operaci. Už z toho je zřejmé, že taková strategie typu „ex post“ je značně neefektivní, neboť zde vzniká riziko zbytečně vynaložených zdrojů.[3], [5], [16]

Vzhledem ke stěžejnímu tématu diplomové práce bude dále statistické regulaci procesu věnován hlavní prostor.

A.3 Statistická regulace procesu

Klasické pojetí statistické regulace procesu („SPC“) předpokládá určit, zda na výrobní proces, respektive regulovanou veličinu, působí vlivy [3], [5]:

- a) Náhodné – takový proces je pod statistickou kontrolou (dle normy ve statisticky zvládnutém stavu), má stálé rozdělení pravděpodobnosti a parametry v požadovaných (neměnných) mezích. Odstranění takového vlivu vyžaduje systémová opatření na taktické/strategické úrovni řízení.
- b) Zvláštní – regulovaná veličina nemá parametry rozdělení trvale na požadovaných úrovních, tedy není pod statistickou kontrolou (dle normy hovoříme o statisticky nezvládnutém procesu). Odstranění takového vlivu je v kompetenci osoby

zodpovědné za proces. „Podmnožinami“ tohoto vlivu jsou dále vlivy sporadické a systémové.

Působení těchto vlivů vyvolává v procesu variabilitu. Konkrétně v souvislosti se SPC zavádíme pojem, který bude vysvětlen dále, tzv. inherentní variability procesu. Jinými slovy lze úvod kapitoly napsat tak, že úkolem SPC je analyzovat vlivy, které inherentní variabilitu vyvolávají, a vytvářet podmínky pro minimalizaci jejich působení. [3], [5]

Jak bylo uvedeno výše, proces je charakterizován určitými výstupními veličinami. V aplikaci statistické regulace se definují tzv. regulované veličiny procesu, kterými mohou být [3], [5]:

- a) znaky jakosti výrobku/služby (například rozměr, pevnost, apod.);
- b) technologické parametry (chemické reakce, technologické zpracování, apod.).

A.3.1 Etapy statistické regulace

Přípravná etapa

V první řadě provedeme volbu znaku jakosti, znak může být jeden nebo může být více znaků. Dále stanovíme kontrolní místa v procesu nejlépe tak, abychom zachytili vznik odchylky. Zvolíme vhodnou délku kontrolního intervalu dle zásad, které jsou uvedeny v literatuře [3], jež je přebírá od dalších autorů:

- kratší interval zvolit na začátku implementace SPC a poté postupně interval prodlužovat;
- zohlednit, jak často se může některý neobvyklý vliv objevit a jak rychle ho chceme zachytit;
- tam, kde se výrobní podmínky často mění, volit krátký kontrolní interval, ve stálých podmínkách je možno volit interval delší.

Výběr logických podskupin

Dále provedeme výběr logických podskupin, jejichž smysl bude vysvětlen dále v textu. Už zde je ale namístě říct, že vhodně zvolený rozsah výběru logických podskupin je důležitý pro správnou aplikaci regulačního diagramu. Nejlogičtější se jeví výběr respektující časové kontinuum výroby. Ovšem je potřeba respektovat i další faktory jako jsou forma produktu (plno/polo/bez tvaré výrobky), stabilita výrobních podmínek (zohledňující právě formu produktu) a počet zdrojů variability (např. počet operátorů, výrobních zdrojů). [3], [5], [16]

Dalším krokem by měl být vhodně zvolený rozsah výběru. Dle již vzpomenuté [3] platí:

- tam, kde jsou náklady na kontrolu větší než náklady spojené s pozdním odhalením působení zvláštní příčiny, pracovat s menším n ;
- tam, kde jsou náklady na kontrolu menší než náklady spojené s pozdním odhalením působení zvláštní příčiny, pracovat s větším n ;
- pro relativně velkou kritickou změnu procesu volit menší n než pro identifikaci relativně malých změn.

Regulace

Následnou analýzou (např. pomocí histogramu) výrobního procesu stanovíme, zda je proces statisticky stabilní. Pokud ne, pak zjišťujeme příčiny takového jevu a řešíme, jak jej stabilizovat.[3], [5], [16]

V první etapě (rovněž jako „etapa zabezpečení statisticky zvládnutého procesu“) uvedeme neustálený proces do původního stavu s využitím regulačního diagramu. Klademe důraz na to, aby se působení zvláštních vlivů neopakovalo. Neustálený proces se projevuje rychlými změnami rozdělení pravděpodobnosti regulované veličiny. Na to je potřeba reagovat i odpovídající změnou podmínek regulace. V této souvislosti může nastat chyba 2. druhu (riziko chybějícího signálu). Ta znamená, že v procesu, který je jinak v nezvládnutém stavu, se hodnota výjimečně objeví uvnitř regulačních mezí. [3], [5], [16]

V druhé etapě (rovněž jako „etapa dlouhodobé regulace procesu“) se proces udržuje v ustáleném stavu, zde s použitím regulačního diagramu stability. Výkyvy nad meze regulačního digramu by měly být v této fázi procesu jen občasné, a pokud jsou, měly by být odstraněny. V této souvislosti může nastat chyba 1. druhu (riziko zbytečného signálu). To znamená, že v procesu, který je jinak ve zvládnutém stavu, se ojediněle vyskytne hodnota mimo meze regulačního diagramu. S tímto nesprávným závěrem jsou spojeny další náklady na (zbytečnou) detekci neexistujícího problému.[3], [5], [16]

A.3.2 Regulační diagramy

Regulační diagram představuje vhodný nástroj pro sledování regulovaných veličin, resp. pro sledování variability procesu v čase. Na vodorovnou osu (osu x) se nanáší časové pořadí kontrol, které mají charakter logických podskupin. Logická podskupina (podmnožina hodnot) znamená výběr určitého množství prvků n z množiny všech hodnot v konkrétní časový okamžik. Výběr je proveden tak, aby se uvnitř podskupiny projevila

variabilita vyvolaná pouze náhodnými příčinami. Projev zvláštní příčiny je minimální v rámci podskupiny, ale maximální mezi podskupinami. Rozdíl hodnot, který nabývají jednotlivé prvky v rámci jedné podskupiny, se nazývá již zmíněná inherentní variabilita procesu.[3], [5], [16]–[18]

Z každé logické podskupiny (n -tice) ve výběru jsou spočítány požadované statistické charakteristiky neboli výběrové ukazatele (výběrový průměr, výběrová směrodatná odchylka, výběrový medián, výběrové rozpětí, počet neshodných jednotek v podskupině, počet neshod na jednotku v podskupině apod.). Ty jsou samostatně vynášeny na svislou osu y . Vodorovnou osou v grafu je dále vyznačena centrální přímka (CL). Ta může být definována jako [3], [5], [16]–[18] :

- odhad z hodnot regulované veličiny, pokud byl proces ve statisticky zvládnutém stavu (pro druhou etapu SPC);
- hodnota jmenovitá, nebo dána technickým předpisem;
- hodnota založená na minulých zkušenostech s výrobním procesem (pro čtvrtou etapu SPC).

Další vodorovné přímky tvoří regulační meze obecně nazvané jako horní (UCL, popř. UWL) a dolní (LCL, popř. LWL) dle vzdálenosti od centrální přímky. U většiny diagramů jsou tyto regulační meze označeny jako cílové (UCL, LCL), výstražné (UWL, LWL) a akční. Regulační meze určují, do jaké oblasti by naměřená data měla spadat v případě ustáleného procesu (tedy když na proces působí náhodné vlivy). Jejich konstrukce je učiněna na základě odhadu inherentní variability procesu. Proto je důležité dodržovat správné dělení dat do logických podskupin, jinak hrozí k nadhodnocení nebo podhodnocení regulačních mezí. Pokud jsou hodnoty v mezích regulace, tedy v cílové oblasti mezi centrální přímkou a mezemi výstražnými, prohlašujeme stabilizaci výrobního procesu za uspokojivou: poloha a proměnlivost jsou stabilní. Pokud se nějaký bod či více bodů vyskytuje mimo regulační meze požadované oblasti nebo body uvnitř této oblasti vytvářejí nenáhodná seskupení, nelze proces považovat za statisticky zvládnutý. Pokud hodnoty padnou do akční oblasti, která leží vně akčních mezí, lze očekávat neobvykle velké posuvy úrovně procesu. Stabilita procesu je tak porušena. V tom případě je nutné vyhledat zvláštní (neobvyklou) příčinu a její působení omezit či plně odstranit. [3], [5], [16]–[18]

Podle způsobu vyhodnocování můžeme hovořit o regulačních diagramech s pamětí a bez paměti. Rozdíl je v tom, jakým způsobem jsou zohledněny předchozí regulované veličiny.[3], [5]

A.3.2.1 Shewhartovy regulační diagramy

Shewhartovy regulační diagramy patří mezi zástupce regulačních diagramů bez paměti. Používají se pro rozhodnutí, zda učinit či neučinit zásah do procesu pouze dle aktuální hodnoty zvolené výběrové charakteristiky. Mezi statistické předpoklady správného použití tohoto typu diagramu patří normalita a nezávislost dat, konstantní střední hodnota a konstantní rozptyl. Hodí se tedy především pro procesy typu A1. Můžeme jejich prostřednictvím zjišťovat kolísání údajů u opakujících se dějů a používat kritéria, která odhalí původ nestability. To znamená, že se hodí tam, kde na proces působí náhodné vnější vlivy s velkými výchylkami. Jak bylo uvedeno výše, data rozdělujeme do podskupin stejného rozsahu, dle stejně měřitelných jednotek. V každé podskupině určujeme aritmetický průměr, medián, variační rozpětí a směrodatnou odchylku. Rozeznáváme dvě situace. Hodnoty základního souboru dat buď jsou, anebo nejsou stanoveny [3], [5], [16]–[19]:

- a) Pokud stanoveny jsou (považujeme je za nultou hodnotu, např. (\bar{x}_0)), musíme je nejprve vyšetřit. Následně zkoumáme, zda se hodnoty z logických podskupin o n počtu pozorování liší od hodnot základního souboru víc, než dovoluje náhodnost.
- b) Pokud stanoveny nejsou, tak v regulačním diagramu zachycujeme výběr dat, kde kolísání hodnot může být způsobeno i jinými než náhodnými vlivy.

Shewhartovy regulační diagramy měřením

Použití diagramů předpokládá spojitou náhodnou veličinu, tzn. normalitu rozdělení dat, konstantní střední hodnotu a rozptyl a nekorelovanost dat. Všechny tyto předpoklady je nutné ověřit pomocí příslušných testů (testy statistických hypotéz, testy normality dat, apod.). Z obrázku č. 1 je patrné, že pro tento druh diagramu se rozlišuje několik typů centrálních přímek. Pokud jde o vzdálenost regulačních mezí, ta je určena směrodatnou odchylkou celkového souboru dat (tedy ne podskupin). Regulační meze jsou postupně ve vzdálenostech $\pm\sigma$, $\pm 2\sigma$, $\pm 3\sigma$. Pravděpodobnost překročení akční meze je 0,3 %. [3], [5], [16]–[19]

Statistika	Základní hodnoty nejsou stanoveny		Základní hodnoty jsou stanoveny	
	Centrální přímka	UCL a LCL	Centrální přímka	UCL a LCL
\bar{X}	$\bar{\bar{X}}$	$\bar{X} \pm A_2 \bar{R}$ nebo $\bar{X} \pm A_3 \bar{s}$	X_0 nebo μ_0	$X_0 \pm A\sigma_0$
R	\bar{R}	$D_4 \bar{R}, D_3 \bar{R}$	R_0 nebo $d_2 \sigma_0$	$D_2 \sigma_0, D_1 \sigma_0$
s	\bar{s}	$B_4 \bar{s}, B_3 \bar{s}$	s_0 nebo $C_4 \sigma_0$	$B_6 \sigma_0, B_5 \sigma_0$
Individuální hodnota X	\bar{X}	$\bar{X} \pm E_2 \bar{R}$	X_0 nebo μ_0	$X_0 \pm 3\sigma_0$
Klouzavé rozpětí R	\bar{R}	$D_4 \bar{R}, D_3 \bar{R}$	R_0 nebo $d_2 \sigma_0$	$D_2 \sigma_0, D_1 \sigma_0$

Obr. 1 Shewartovy regulační diagramy měřením [3]

Diagram se konstruuje pro dvojici parametrů, kdy jeden charakterizuje polohu a druhý měnlivost procesu. Nejčastěji se používají kombinace diagramů průměr-variační rozpětí a průměr-směrodatná odchylka.[3], [5], [7], [16]

- a) Regulační diagramy pro výběrový průměr a výběrové rozpětí (\bar{x}, R)

Tato dvojice regulačních diagramů je vhodná, pokud máme množinu dat s menším rozsahem výběrů (do max. $n = 10$ jednotek). Pokud výběr obsahuje minimálně 4 jednotky, je možné tento regulační diagram sestavit i pro data, která nemají normální rozdělení.[7]

- b) Regulační diagramy pro výběrový průměr a směrodatnou odchylku (\bar{x}, s)

Tato kombinace regulačních diagramů je použitelná jak pro malé výběry, tak i pro realizaci výběrů většího rozsahu ($n > 10$).[7]

- c) Regulační diagram pro individuální hodnoty (X)

Regulační diagram pro individuální hodnoty je vysoce citlivý na odchylky rozdělení dat od normálního, a je tedy žádoucí podrobit regulovaná data některému z testů na normalitu.[7]

Shewartovy regulační diagramy srovnáváním

Shewartovy regulační diagramy srovnáváním se použijí, pokud je regulovaná veličina v čase diskretní a náhodná. Hodí se tedy pro zaznamenání ne/přítomnosti určitého znaku a počítá se, kolik jednotek znak ne/vyazuje. Stejně jako u Shewartova regulačního diagramu měřením i zde rozlišujeme různé typy centrálních přímek, jak je vidět na obrázku č. 2. Pro binomické rozdělení konstruueme diagram pro podíl (p), resp. počet (np) neshodných jednotek. Pro Poissonovo rozdělení konstruueme diagram počet neshod (c) a počet neshod na jednotku (u). [3], [5], [8], [16], [17]

Statistika	Základní hodnoty nejsou stanoveny		Základní hodnoty jsou stanoveny	
	Centrální přímka	3σ regulační meze	Centrální přímka	3σ regulační meze
p	\bar{p}	$\bar{p} \pm 3 \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$	n_0	$p_0 \pm 3 \sqrt{\frac{p_0(1-p_0)}{n}}$
np	$n\bar{p}$	$n\bar{p} \pm 3\sqrt{n\bar{p}(1-\bar{p})}$	np_0	$np_0 \pm 3\sqrt{np_0(1-p_0)}$
c	\bar{c}	$\bar{c} \pm 3\sqrt{\bar{c}}$	c_0	$c_0 \pm 3\sqrt{c_0}$
u	\bar{u}	$\bar{u} \pm 3 \sqrt{\frac{\bar{u}}{n}}$	u_0	$u_0 \pm 3 \sqrt{\frac{u_0}{n}}$

Obr. 2 Shewartovy regulační diagramy srovnáváním [3]

A.3.2.2 Regulační diagramy pro aritmetický průměr s výstražnými mezemi

Oproti Shewartovu regulačnímu diagramu je regulační diagram pro aritmetický průměr s výstražnými mezemi citlivější na pozvolné změny. Předpokladem je normální rozdělení dat. Jsou založeny na výpočtu aritmetického průměru μ_0 z výběru, který tvoří centrální přímku. Pokud jde o regulační meze, tak meze výstražné a akční jsou dány postupně vzorci č. 8 a 9 [3], [16], [17]:

$$\mu_0 \pm \frac{B_2 \cdot \sigma}{\sqrt{\sigma}} \quad (8)$$

$$\mu_0 \pm \frac{B_1 \cdot \sigma}{\sqrt{\sigma}} \quad (9)$$

A.3.2.3 Přejímací regulační diagramy

Přejímací regulační diagramy jsou založené na Shewartově diagramu, ale zaměřené na hodnocení, zdali se ukazatel nachází v tolerančním poli, tedy toho, zda je proces vyhovující, nebo ne. Proces musí být statisticky zvládnutý především z hlediska měnlivosti uvnitř podskupiny. Předpokládá se, že příčiny vyvolávající posuny v úrovních procesu jsou dostatečně malé a bylo by příliš nákladné je regulovat. Dále je tento diagram vhodný pro procesy typu C. [3], [16], [17]

Přejímací regulační diagram může být navržen dvěma způsoby, dle kterých se liší výpočet regulačních mezí. Pro způsob konstrukce diagramu, který je nejrozšířenější, se vychází ze zadání hodnot maximálně přípustného a nepřípustného podílu neshodných jednotek a rizika chyby prvního a druhého druhu. Z těchto vstupních parametrů se vypočte velikost logické podskupiny. Pokud střední hodnota procesu leží v oblasti nevyhovujících procesů, tak to znamená, že produkuje nepřijatelný podíl neshodných jednotek a do procesu je potřeba zasáhnout.[3]

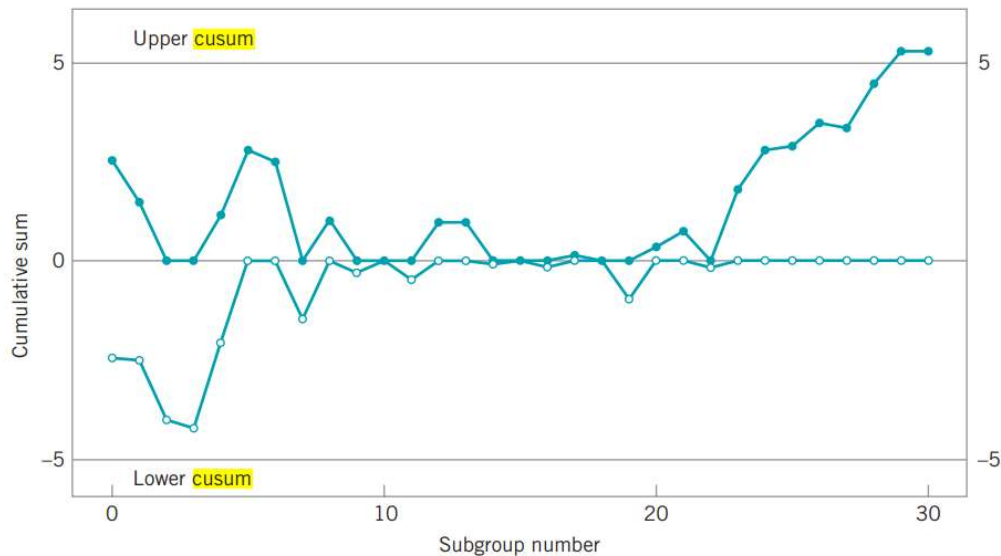


Obr. 3 Přejímací regulační diagram [17]

A.3.2.4 CUSUM regulační diagramy

Metoda CUSUM je vhodná pro detekci systematických vnějších vlivů. V takovém případě totiž selhává použití Shewartových diagramů, jelikož takové vlivy nemohou být odhaleny pomocí jednotlivých izolovaných výběrů. Naopak je potřeba metody, která akumuluje informaci obsaženou ve všech za sebou jdoucích výběrech. Tou je právě metoda CUSUM (z anglického „kumulovaný součet“). Spočívá v procesu odečítání předem dané hodnoty (cílové, preferované) od každého pozorování v řadě a postupného načítání těchto rozdílů. Takový proces je ve statistice nazýván náhodné kráčení. Regulační diagramy CUSUM interpretují data uspořádaná ve smysluplné posloupnosti. Na ose x se zaznamenává pořadí výběru a na ose y se vynášejí hodnoty testovacího kritéria S_i . [3], [15], [20]

$$S_i = \sum_{j=1}^i \bar{X}_j - \mu_0 = S_{i-1} + (\bar{X}_j - \mu_0), S_0 = 0 \quad (10)$$



Obr. 4 Příklad CUSUM regulačního diagramu [20]

Pokud je proces udržován na cílové hodnotě μ_0 , pak body v diagramu zachovávají směr rovnoběžný s osou x. Jestliže střední hodnota procesu roste, nebo klesá a ještě se nestabilizovala, pak body v diagramu tvoří křivku viditelně se zakřivující nahoru nebo dolů. Všechny výše uvedené jevy je možné pozorovat na obrázku č. 4. Regulační meze jsou dány druhem CUSUM regulačního diagramu. Ve srovnání s Shewartovými regulačními diagramy detekuje tato metoda změny významné velikosti až 3x rychleji.[3], [15], [20]

A.3.2.5 EWMA diagramy

EWMA diagramy neboli diagramy exponenciálně vážených klouzavých průměrů (z anglického „exponentially weighted moving average“). Obdobně jako diagram CUSUM umožňuje rychlejší detekci malých až středně velkých posunů. V tomto diagramu se průměr procesu určuje jako exponenciálně vážený klouzavý průměr všech předcházejících průměrů podskupin. EWMA přiřazuje podskupinám váhu klesající geometricky tak, že nejnovější podskupiny mají nejvyšší váhu. [3], [15], [20]

$$EWMA_i = (1 - \lambda)^i \cdot \mu_0 + \lambda \cdot \sum_{j=1}^i (1 - \lambda)^{i-j} \cdot \bar{x}_j \text{ pro } 0 < \lambda < 1, j = 1, 2, \dots, k \quad (11)$$

Čím je parametr λ větší, tím je menší vliv starších výběrových průměrů, a čím více se blíží hodnota parametru λ k 1, tím více se blíží EWMA diagram klasickému Shewartovu diagramu bez paměti. Naopak pokud λ se blíží k nule, tím je průběh diagramu plynulejší a méně citlivý na rychlou změnu. Navíc je tento modul výchozí pro konstrukci dynamického diagramu EWMA. Tyto diagramy jsou rovněž robustní vůči nenormálnímu

rozdělení dat. I zde je centrální přímkou tvořena aritmetickým průměrem μ_0 . Vztahy pro výpočet regulačních mezí jsou dány vybranou metodou EWMA diagramu. [3], [15], [20]

Literatura (Pokročilé metody) rozlišuje mezi diagramem EWMA pro regulaci střední hodnoty a metodou FIR EWMA. [3]

A.3.3 Volba typu regulačního diagramu

Literatura [3], [5], [15], [16] obecně hovoří o několika vstupních podmínkách pro tvorbu regulačních diagramů.

Paní Tůmová [16] zmiňuje dva možné způsoby regulace, ze kterých se vybírá dle požadavků na přesnost, přesnost a účinnost mezioperační kontroly. Jedná se o regulaci měření, kdy sledovaný znak je vyjádřen spojitou kvantitativní veličinou, a srovnáváním, kdy sledovaný znak je vyjádřen veličinou kvalitativní.

Kvantitativními veličinami u regulace měření se rozumí elementy popisné statistiky polohy (průměr, kvartily, medián) a měnlivosti (variační rozpětí, směrodatná odchylka). Sestrojují se dva regulační diagramy právě pro charakteristiky polohy a měnlivosti. (Volba diagramů závisí na rozsahu výběru v podskupině.) U kvalitativních veličin zjišťujeme ne/přítomnost určité vlastnosti, kterou zachycujeme absolutní a relativní četností. Sestrojuje se jeden diagram, a to buď p (resp. np), nebo c (resp. u) diagram podle toho, zda je rozsah n konstantní u neshod a neshodných jednotek.[16]

Paní Noskiewičová [3] píše, že „modernější“ pojetí SPC počítá s dělením procesů do skupin, jak je znázorněno na obrázku č. 5. Rozdělení do skupin je učiněno dle změn jejich střední hodnoty a rozptylu nebo dle tvaru rozdělení, a to v průběhu času (ať okamžité, nebo výsledné hodnoty).

Model	Okamžité rozdělení	Parametr polohy	Parametr variability	Výsledné rozdělení
A1	normální	konstantní	konstantní	normální
A2	jednovrcholové	konstantní	konstantní	jednovrcholové
B	normální	konstantní	systematické a náhodné změny	jednovrcholové
C1	normální	náhodné změny	konstantní	normální
C2	normální	náhodné změny	konstantní	jednovrcholové
C3	libovolné	systematické změny	konstantní	libovolné
C4	libovolné	systematické a náhodné změny	konstantní	libovolné
D	libovolné	systematické a náhodné změny	systematické a náhodné změny	libovolné

Obr. 5 Přehled typu procesů [3]

S ohledem na zvolený proces je mj. vybrán regulační diagram, jak lze vidět na obrázku č. 6.[3]

Nesplněný předpoklad	Typ procesu	Metoda SPC
Žádný	A1	Klasické Shewhartovy regulační diagramy
Normalita dat	A2	Regulační diagramy s nesymetrickými mezemi
Neměnnost parametrů rozdělení	B, C1 – C4, D	Modifikované regulační diagramy Přijímací regulační diagramy Regresní regulační diagramy Regulační diagramy s rozšířenými mezemi
Nezávislost dat	Stacionární, nestacionární proces	Diagramy ARIMA Diagramy pro rezidua EWMA modelu Dynamické diagramy EWMA Diagramy pro skupinové průměry
Vysoký stupeň opakovatelnosti procesu	Malosériová výroba	Cílové diagramy Standardizované diagramy Q-diagramy Hillierova metoda Předregulace
Velké změny procesu	Vysoce přesná výroba	Diagramy CUSUM Diagramy EWMA
Sledování pouze jednoho znaku	Simultánní sledování několika znaků	Hotellingovy diagramy T^2 Vícerozměrné CUSUM Vícerozměrné EWMA
	Procesy s vysokou mírou způsobilosti	Diagramy CCC, CCC-r CCC CUSUM, CCC-r CUSUM CCC-EWMA, CCC-r EWMA

Obr. 6 Volba regulačního diagramu dle typu procesu [3]

Proces typu A

Proces typu A předpokládá konstantní střední hodnotu a rozptyl v průběhu času. Okamžitý tvar rozdělení může být normální nebo jednovrcholový. Platí zde, že výsledné rozdělení je totožné s tím okamžitým.[3]

Proces typu B

Proces typu B předpokládá konstantní střední hodnotu a variabilní rozptyl v průběhu času. Okamžité tvary rozdělení jsou normální, ale z důvodu jejich různých rozptylů je jejich výsledné rozdělení jednovrcholové a nenormální. Variabilní rozptyl může být dán fluktuací materiálu.[3]

Proces typu C

Proces typu C předpokládá proměnlivou střední hodnotu, ale konstantní rozptyl. Typ rozdělení závisí na tom, zda je střední hodnota ovlivněna vlivy systematickými nebo náhodnými. Pokud působí vlivy náhodné, je tvar okamžitého rozdělení normální a stejně tomu tak je i ve výsledném rozdělení. Při působení systematických vlivů neřešíme tvar okamžitého rozdělení, to výsledné rozdělení je jednovrcholové nenormální.[3]

Proces typu D

Takové procesy jsou charakteristické variabilitou rozptylu, střední hodnoty a tvaru rozdělení.[3]

A.4 Nenormalita dat

Jak bylo řečeno dříve, některé aplikace statistické regulace (ale i třeba hodnocení způsobilosti procesu) vyžadují normální rozdělení dat. Avšak mohou nastat situace, kdy podmínka normality nemusí být zcela splněna. Pokud by byla nenormálně rozdělená data vyhodnocena stejným způsobem jako normálně rozdělená data, dojde ke zkreslení výsledku. Ověření normality dat je možné provést dvěma metodami, a to buď grafickou, nebo početní. [3], [10], [13], [15], [21]–[23]

A.4.1 Grafické testy normality

Grafické metody jsou vhodné spíše pro předběžné posouzení normality.[3], [13], [21]

Histogram

Histogram je možným, ale méně přesným nástrojem pro vyhodnocení normality. Pro malý počet hodnot nelze spolehlivě prohlásit podobnost takového histogramu s křivkou hustoty normálního rozdělení.[3], [13], [21]

Q-Q graf

Princip Q-Q grafu spočívá v tom, že na jednu osu jsou nanášeny **kvantily** normovaného normálního rozdělení u_i (popř. hypotetického normálního rozdělení $N(\mu, \sigma^2)$) a na druhou osu kvantily zkoumaného souboru. V případě normálního rozdělení leží body na přímce. [3], [13], [21]

P-P graf

P-P graf je obdobou předchozí grafické metodě s tím rozdílem, že místo kvantilů se na jednu osu vynáší **hodnota distribuční funkce teoretického normálního rozdělení** Fx_i . Na druhou osu se pak vynáší teoretické distribuční funkce. P-P grafy lépe odhalují odchylky od normality v oblasti kolem střední hodnoty. Q-Q grafy zase lépe identifikují odlišný tvar konců rozdělení. [3], [13], [21]

A.4.2 Početní testy normality

Grafické testy popsané v předchozí kapitole jsou vhodné pro předběžné posouzení normálního rozdělení. Na druhé straně přesnější (početní) metody testování skutečně zkoumají hypotézu H_0 , zda jsou data z normálního rozdělení. Pokud hypotézu H_0 nezamítáme, považujeme předpoklad normality za splněný. Metodu testováním lze rozdělit do několika skupin, a to dle charakteru (principu) prováděných testů. Jsou to testy regresní, testy založené na momentech a testy založené na empirických distribučních funkcích. [3], [13], [21]

A.4.2.1 Testy založené na empirických distribučních funkcích (tzv. testy dobré shody)

Všechny testy z této skupiny jsou založeny na porovnání empirické distribuční funkce s teoretickou distribuční funkcí normálního rozdělení. [3], [13], [21]

Kolmogorovův-Smirnovův test

Slovně lze tuto statistiku definovat tak, že měříme maximální vzdálenost mezi empirickou a teoretickou distribuční funkcí. Pokud x_i jsou pozorování uspořádaná vzestupně, pak je testová statistika dána [3], [13], [21]:

$$D_n = \max \left\{ \left| F(x_i) - \frac{i-1}{n} \right|, \left| F(x_i) - \frac{i}{n} \right| \right\} \quad (12)$$

H_0 zamítáme, pokud $D_n > D_\alpha$.

Anderson-Darlingův test / Cramérův-von Misesův test

Jako u předchozího testu i zde měříme vzdálenost mezi empirickou a teoretickou distribuční funkcí, v tomto případě se však jedná o čtvercovou vzdálenost. Anderson-Darlingův test je navíc doplněn o váhy, viz vztah [3], [13], [21]:

$$A^2 = -n - \sum_{i=1}^n \frac{2i-1}{n} (\ln z_i + \ln(1 - z_{n+1-i})) \quad (13)$$

Pokud hodnota testové statistiky A^2 překročí kritickou hodnotu $D_{1-\alpha}$, zamítáme hypotézu H_0 .

Test Chí-kvadrát

Poprvé byl test Chí-kvadrát prezentován v roce 1900 panem Pearsonem. Dnes je již překonaný „silnějšími“ testy a v této práci je uveden pouze pro úplnost. [3], [13], [21]

A.4.3 Práce s nenormálním rozdělením a jeho příčiny

Pokud jsme data řádně otestovali a eventualita nenormálního rozdělení analyzovaných dat nastala, je tu několik možností, jak se s takovým rozdělením vypořádat. V první řadě lze použít testy, které jsou natolik robustní, že samotný fakt nenormálního rozdělení nezkrusí výsledek. Takovými jsou např. t -test (jedno i dvouvýběrový a párový test), test analýzy rozptylu (ANOVA test) a dále regresní testování. V druhé řadě je možné data transformovat a získat tak přibližně normální data, přičemž pro vyhodnocení dat je nutné provést zpětnou transformaci. Obecně je však lepší nejprve hledat příčiny nenormálního rozdělení dat a ty odstranit, protože při transformaci se do určité míry skryje fyzikální význam dat a nelze do nich jednoduše nahlížet. [10], [22], [23]

Možnými příčinami nenormálního rozdělení dat mohou být chyba v jejich zápisu, chybějící data, dva sety v jednom souboru dat, odlehlé a extrémní hodnoty. Chyba v zápisu

se eliminuje jen těžko. Chybějící data mohou být vyřešena vyloučením neúplně popsanych objektů – tedy odebráním dat, vyloučením neúplně popsanych párů, nahrazením chybějící hodnoty aritmetickým průměrem okolních hodnot a regresní aproximací. Pokud by se nacházela dvě měření v jednom souboru dat, lze sestavit XY graf k detekci takové příčiny nenormálního rozdělení. Odlehlé a extrémní hodnoty lze vyšetřit kvartilovou metodou a Dean-Dixonovým, popř. Grubbsovým testem. Platí, že specifitější může být detekce odlehlých hodnot, protože na rozdíl od hodnot extrémních, tyto se nacházejí blízko základního souboru hodnot a nelze je určit pouhým subjektivním posouzením. [10], [22], [23]

A.4.3.1 Detekce odlehlých hodnot

Pro všechny níže uvedené postupy platí, že nejprve je potřeba hodnotu podezřelou na odlehlost odebrat ze souboru dat. [10], [22], [23]

Kvartilová metoda

Pokud podezřelá odlehlá hodnota maxima, resp. minima leží mimo níže uvedené intervaly, lze ji prohlásit za odlehlou. [10], [22], [23]

$$HVH = \bar{x}_{75} + 1,5 \cdot (\bar{x}_{75} - \bar{x}_{25}) \quad (14)$$

$$DVH = \bar{x}_{25} - 1,5 \cdot (\bar{x}_{75} - \bar{x}_{25}) \quad (15)$$

Grubbsův test

Vypočtenou hodnotu testovacího kritéria porovnáme s kritickou hodnotou T_α na hladině významnosti. Je-li $T_1 \geq T_\alpha$, resp. $T_n \geq T_\alpha$, zamítáme H_0 , hodnotu tedy považujeme za odlehlou a můžeme ji vyloučit ze souboru. [10], [22], [23]

$$T_1 = \frac{\bar{x} - x_1}{s} \quad (16)$$

$$T_n = \frac{\bar{x}_n - \bar{x}}{s} \quad (17)$$

Dean-Dixon test

Stejně jako u předchozího testu seřadíme hodnoty vzestupně a vypočteme testovací kritérium pro první, respektive poslední hodnotu ze souboru. Pokud $Q_1 \geq Q_\alpha$, resp. $Q_n \geq Q_\alpha$, pak opět můžeme hodnotu vyloučit ze souboru. [10], [22], [23]

$$Q_1 = \frac{x_2 - x_1}{x_{max} - x_{min}} \quad (18)$$

$$Q_n = \frac{x_n - x_{n-1}}{x_{max} - x_{min}} \quad (19)$$

Jsou však případy, kdy data nemohou mít z fyzikální podstaty věci normální rozdělení, a transformaci se tak nevyhneme. Samotná transformace vede ke stabilizaci rozptylu, zesymetričtění rozdělení a je robustní k odlehlým bodům. [10], [22], [23]

A.4.3.2 Transformace dat

Pojem „transformace“ lze v tomto případě zaměnit za slovo „přepočítání“. Data s nenormálním rozdělením jsou přepočítána (tedy transformována) speciálně zvolenou funkcí tak, aby výsledná data měla rozdělení normální. Tato funkce musí být nelineární, aby nedošlo pouze k posunu dat, a monotónní, aby nedošlo ke změně pořadí velikosti. [3], [15]

Samotné provedení transformace nezaručí normalitu transformovaných dat, předpoklad normality je potřeba následně znovu otestovat. Nicméně při úspěšné transformaci lze na datech dále provádět všechny testy tak, jako kdyby měla nativně rozdělení normální. [3], [15]

Pro transformaci dat lze využít následující metody. Jako nezákladnější metoda je v literatuře [3], [15] uvedena mocninná transformace. Za méně účinnou bývá popisována Box-Coxova transformace a navíc použitelná jen pro kladné hodnoty. Naproti tomu Johnsonova transformace se jeví jako účinnější a lze ji použít pro nulové i záporné hodnoty.

Mocninná transformace

$$y = \begin{cases} x^\lambda & \lambda > 0 \\ -x^\lambda & \lambda < 0 \\ \ln x & \lambda = 0 \end{cases} \quad (20)$$

Box-Coxova transformace

$$y = \begin{cases} \frac{x_i^\lambda - 1}{\lambda \cdot g \cdot \ln x} & \lambda \neq 0 \\ g \cdot \ln x & \lambda = 0 \end{cases} \quad (21)$$

$$g = \left(\prod_{i=1}^n x_i \right)^{\frac{1}{n}} \quad (22)$$

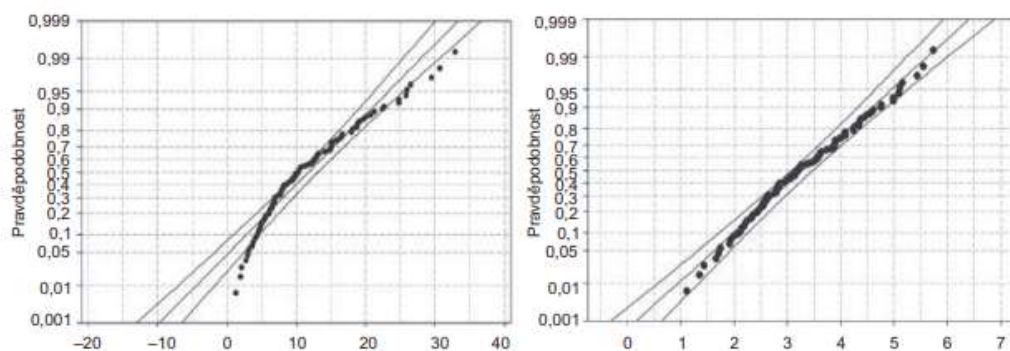
Johnsonova transformace

$$y = \begin{cases} \gamma + \eta \ln\left(\frac{x_i - \varepsilon}{\lambda + \varepsilon - x_i}\right) & \text{pro SB} \\ \gamma + \eta \ln\left(\frac{x_i - \varepsilon}{\lambda}\right) & \text{pro SL} \\ \gamma + \eta \sinh^{-1}\left(\frac{x_i - \varepsilon}{\lambda}\right) & \text{pro SU} \end{cases} \quad \begin{cases} \eta, \lambda > 0, -\infty < y, \varepsilon < \infty, \varepsilon < x_i < \varepsilon + \lambda \\ \eta, \lambda > 0, -\infty < y, \varepsilon < \infty, \varepsilon < x_i \\ \eta, \lambda > 0, -\infty < y, \varepsilon, x_i < \infty \end{cases} \quad (23)$$

λ se volí tak, aby minimalizovala směrodatnou odchylku náhodné veličiny (obvykle v rozsahu -5 až 5).

Transformace dle níže uvedených křivek

Na obrázku č. 7 je znázorněn princip transformace.



Obr. 7 Příklad transformace dat [3]

Transformovaná data je nutné po analýze zpětně transformovat funkcí inverzní k transformační, abychom dostali původní data, protože právě jejich analýzu požadujeme. [3], [15]

A.5 Statistický software

Před výběrem softwaru pro statistickou analýzu je nutné zvážit několik faktorů.

Datový typ

Každý software nabízí různé rozsahy funkcí a použitých algoritmů, které uživatel nemusí všechny potřebovat. Zvolený software by měl provádět přesný druh analýzy dat, kterou uživatel potřebuje. Například při práci s velkým množstvím dat by bylo efektivnější investovat do nástroje určeného k rychlému zpracování velkých datových sad.[24]–[28]

Uživatelské prostředí

Zde je potřeba si uvědomit, kdo bude software používat a jaká je jeho úroveň zkušeností se statistikami. Pokročilejší uživatelé například snadno zvládnou programovací jazyky a složité rovnice. Začínající uživatelé se však mohou cítit zahlceni používáním softwaru s podobným formátem a upřednostní software s nabídkovým rozhraním. [24]–[28]

Grafické výstupy

V tomto směru je potřeba rozmyslet, zda požadujeme, aby software pro statistickou analýzu okamžitě vytvářel vizuální grafiku a zda má být tato vizualizace interaktivní nebo statická. Dále je potřeba uvážit, jestli bude vizualizace odesílána do jiného programu. V takovém případě musí být exportovaná grafika ve formátu, který běžně používáme. [24]–[28]

A.5.1 Tibco statistica

Tibco Statistica je flexibilní analytický systém, který umožňuje uživatelům vytvářet analytické workflow. Tato technologie také pomáhá interaktivně zkoumat a vizualizovat problémy, vytvářet a zavádět statistické, prediktivní, optimalizační a textové analýzy. Statistica nabízená v rámci platformy Tibco Insight umožňuje uživatelům využívat robustní nástroje pro modelování a ověřování a rovněž pro strojové a hloubkové učení. To v důsledku přináší inteligentnější a správně načasované rozhodování. Software má opakovaně použitelné šablony, které lze pohodlně aplikovat, monitorovat a řídit ve výrobě, což pomáhá uživatelům k porozumění problémů a studiu analytických přístupů. [28]

Výhody:

- prediktivní analýza;
- přístup k datům a příprava;
- nativní distribuovaná analytická architektura (NDAA);

- zabezpečení a správa analytiky;
- otevřený, flexibilní, rozšiřitelný.

Nevýhody

- TIBCO Data Science určený pro pokročilé uživatele;
- obtížné připojení k jiným databázím.

A.5.2 IBM SPSS

SPSS je považován za uživatelsky přívětivý a je jedním z nejpoužívanějších statistických programů. Díky snadno použitelnému rozhraní je vhodný pro začátečníky, ale je i účinnou platformou pro odborníky. V novějších verzích je software vyvíjen společností IBM. Vývoj v posledních letech vedl k tomu, že obraz SPSS byl v očích odborné komunity poněkud degradován, jelikož je software pověstný tím, že jej používají uživatelé, kteří se v dané problematice dostatečně neorientují. IBM přistupuje ke statistické analýze modulárně, což v praxi znamená, že si uživatel může vybrat, co potřebuje. Modul statistik je dodáván s desítkami funkcí, jako jsou rozhodovací stromy, regrese, prognózy a schopnost odhalit chybějící hodnoty. Existují také desítky dalších produktů SPSS, které lze kombinovat a vytvořit přizpůsobenou platformu pro statistické analýzy.[27]

Výhody:

- snadná orientace;
- lze rozšířit pomocí komerčních modulů;
- rozsáhlá literatura.

Nevýhody:

- stabilita trpěla pouze ročním aktualizacím cyklem;
- navzdory syntaxi a skriptovému jazyku je automatizace a integrace do jiných aplikací obtížnější než jiná softwarová řešení.

A.5.3 Stata

Stata je komplexní statistické řešení určené pro různé oblasti, včetně lékařství, ekonomiky, financí, behaviorálních věd a marketingu. Je vhodný pro manipulaci s daty, průzkum, vizualizaci a statistickou analýzu. Software má nejen grafické uživatelské rozhraní, ale i strukturu příkazového řádku, a je tak přístupný jak uživatelům se znalostmi kódování, tak i bez nich. Díky své univerzálnosti a široké škále použití se jedná o oblíbený software pro statistickou analýzu typu all-in-one. Kromě rozšířené regrese,

generalizovaných a vybraných modelů nabízí Stata také tisíce vestavěných bayesovských modelů a pomáhá uživatelům porozumět časově invariantním funkcím. Vyvolává také balíčky Pythonu (Pandas, TensorFlow atd.) pro předávání dat a vizualizaci. Díky těmto integracím Pythonu je strojové učení snazší než kdy dříve, a dokonce je možné přímo v Pythonu spouštět statistické analýzy. Se Stata je možné nastavit stovky tisíc proměnných, zpracovávat textová a binární data na jedno kliknutí, porovnávat, slučovat a připojovat datové sady, spravovat mnoho datových sad současně, a dokonce exportovat nebo importovat data. [26]

Výhody:

- široká škála funkcí - ve STATA najdete téměř každou zavedenou statistickou metodu;
- snadno přístupné prostřednictvím GUI;
- lze automatizovat;
- kompatibilní se staršími verzemi;
- dobrá podpora komunity prostřednictvím rozsáhlé dostupné literatury.

Nevýhody:

- pomalé z hlediska začlenění nových metod (dlouhá doba mezi aktualizacemi verzí);
- integrace s jiným softwarem je problematická;
- omezeno na sadu dat, která je otevřena současně.

A.5.4 SAS

SAS/STAT nabízí komplexní nástroje statistické analýzy, které splňují požadavky na případné rozšiřování. Obsahuje techniky pro malé datové sady a moderní metody pro analýzu dat s chybějícími hodnotami, stejně jako vysoce výkonné nástroje pro statistické modelování pro velké datové úlohy. SAS Institute nabízí profesionální statistický software, který se běžně používá v biometrii, klinickém výzkumu a v bankovním sektoru. Uživatelé mohou data okamžitě vizualizovat, určovat odlehle hodnoty, zaznamenávat trendy, předpovídat budoucí výsledky a identifikovat korelace. Software má dále prediktivní modelování založené na analyzovaných datech pro výpočet pravděpodobných a možných výsledků spolu s vestavěnými postupy statistické analýzy, které vyhovují specializovaným potřebám. SAS je výkonný a velmi stabilní nástroj, který je velmi využíván velkými organizacemi a stal se standardem pro mnoho analýz ve farmaceutickém sektoru. Software se skládá z různých modulů, z nichž některé sledují zcela odlišné provozní koncepce.

Ve srovnání s jinými komerčními konkurenty je SAS jedním z nejdražších řešení (což je částečně dáno jeho zaměřením na větší společnosti a organizace).[25]

Výhody:

- rychlá integrace nových statistických metod, velmi stabilní a spolehlivé rutiny;
- velmi dobrá dokumentace a časově neomezená profesionální podpora;
- k dispozici řada modulů a rozhraní a také vlastní software Business Intelligence (za poplatek);
- vhodný pro zpracování velkých datových sad;
- otevřený vývoj modelu založený na kódu;
- možnost rozsáhlého interního školení;
- distribuované analytické zpracování v paměti, popisné a prediktivní modelování; průzkum a objevování vizuálních dat.

Nevýhody:

- různé, částečně komplikované (ale výkonné) programovací jazyky;
- částečně zastaralé rozhraní, GUI volitelně;
- licenční model a jeho cena (roční licence pro SAS® Analytics Pro začíná přibližně na 7500 €).

A.5.5 Matlab

Systém MATLAB (z anglického Matrix Laboratory) je interaktivní programové prostředí a skriptovací programovací jazyk určený pro vědecké a inženýrské výpočty a vizualizaci dat. Program je vyvíjen od roku 1984 společností The Mathworks, Inc. v USA. Původně byl jako jazyk určen pro matematické účely, ale postupem času byly přidány nové funkce a rozšíření, rozrostl se různými směry a dnes je využitelný v široké paletě aplikací. Aplikace MATLAB umožňují iterační úkoly provádět interaktivně. Aplikace pak generují kód MATLAB potřebný k programové reprodukci práce, kterou uživatel provedl interaktivně. Program disponuje předem připravenou skupinou funkcí pro identifikaci odlehklých hodnot, chybějících dat a šumu. [24]

Výhody:

- datové typy a možnosti předzpracování určené pro technická a vědecká data;
- interaktivní a vysoce přizpůsobitelné vizualizace dat;

- tisíce přednastavených funkcí pro statistickou analýzu, strojové učení a zpracování signálu;
- rozsáhlá a odborně napsaná podpůrná dokumentace;
- zrychlený výkon s jednoduchými změnami kódu a dalším hardwarem;
- rozšířená analýza na velká data bez velkých změn kódu.

A.6 Normy

ČSN ISO 11462 poskytují směrnice pro uplatňování systému statistické regulace procesu. Tyto směrnice uvádějí do procesu takové prvky, které jsou zaměřeny na zvyšování účinnosti a způsobilosti produkce a na snižování prostojů a nákladů. To je mj. naplněno tím, že jsou prohlubovány znalosti o procesu, proces je řízen tak, aby se choval požadovaným způsobem, a je snižováno kolísání parametrů konečného produktu.

ČSN ISO 11462-1 – Směrnice pro uplatňování statistické regulace procesu (SPC) - Část 1: Prvky SPC

Tato část normy poskytuje prvky, které povedou organizaci při plánování, rozvoji, realizování nebo hodnocení systému statistické regulace procesů. Dále rozšiřuje definici procesu s cílem integrovat tradiční definice statistické a algoritmické regulace procesu a metod regulace založených na modelech. Dále rozšiřuje definici a použití termínu parametr ve smyslu procesu a produktu, přičemž produktem se v procesu rozumí produkce nebo finální produkt.

ČSN ISO 11462-2 – Směrnice pro uplatňování statistické regulace procesu (SPC) - Část 2: Katalog nástrojů a postupů

Tato část normy představuje katalog nástrojů a postupů na pomoc organizacím při plánování, zavádění a hodnocení efektivního systému statistické regulace procesu.

ČSN 01 0266 – Zvláštní typy statistické regulace. Metoda kumulovaných součtů

Tato norma stanoví zásady použití zvláštního typu statistické regulace, metody kumulovaných součtů. Tato metoda je vhodná pro statistickou regulaci výrobního procesu, které již byly dříve uvedeny do stabilního stavu. Vyznačuje se vysokou citlivostí na malé změny ve stabilitě procesu a názorností.

ČSN ISO 7870-2 (010272) – Regulační diagramy - Část 2: Shewhartovy regulační diagramy

Tato část normy poskytuje návod a pochopení přístupu ke statistické regulaci založeného na Shewartově regulačních diagramech. Jsou zde představeny doplňující metody konzistentní se Shewartovým přístupem, jako je použití varovných mezí, analýza trendů a způsobilost procesu.

ČSN ISO 7870-4 (010272) – Regulační diagramy - Část 4: Regulační diagramy CUSUM

Tato část normy stanovuje statistické postupy pro určení schémat založených na kumulativním součtu pro řízení procesu a kvality pomocí měřitelných a diskrétních dat.

Popisuje obecně použitelné metody pro rozhodování založené na metodě kumulovaných součtů.

ČSN ISO 7870-6 (010272) – Regulační diagramy - Část 6: Regulační diagramy EWMA

Tato část normy zahrnuje regulační diagramy EWMA jako metodu statistické regulace procesu pro detekci malých posunů střední hodnoty.

ČSN ISO 2854 – Statistická interpretace údajů – Síla testů středních hodnot a rozptylu

Tato norma stanovuje postupy pro odhad střední hodnoty nebo rozptylu základních souborů dat. Dále vyšetření určitých hypotéz souvisejících s hodnotami těchto parametrů pomocí výběrů.

ČSN ISO 3534-1 – Statistika – Slovník a značky – Část 1: Obecné statistické termíny a termíny používané v pravděpodobnosti.

Tato část normy definuje obecné statistické termíny a termíny používané v pravděpodobnosti.

ČSN ISO 3534-2 – Statistika - Slovník a značky – Část 2: Aplikovaná statistika

Tato část normy definuje termíny z oblasti aplikované statistiky a uvádí je do systému pojmů v souladu se zvyklostí normalizace terminologie v ISO.

ČSN ISO 5479 – Statistická interpretace údajů – Testy odchýlení od normálního rozdělení

Tato mezinárodní norma obsahuje návod na použití metod a testů k rozhodnutí, zda je či není možno zamítnout hypotézu o normálním rozdělení za předpokladu, že pozorování jsou nezávislá.

B. Praktická část

K praktické části práce byla dodána data z výrobního procesu. Sledovanou proměnnou je teplota skelného přechodu T_g . Tato teplota by měla být 110 °C, přičemž předpokládáme výrobní meze ± 5 °C. Pro její vyhodnocení sledujeme parametry „operátor“ a „velikost dávky“.

Zadání práce stanovuje, že vyhodnocení je nutné udělat minimálně dvěma statistickými softwary. Pro první (základní) vyhodnocení bude použit Excel z balíčku Microsoft Office. Tato část práce bude vypracována i s ohledem na požadavky následné analýzy v softwaru Statistica.

B.1 Práce v softwaru Excel

Prvotní rozřídění dat bylo provedeno dle jednotlivých operátorů a dle samotných velikostí dávek. Zároveň byl vytvořen graf všech operátorů z hodnot T_g v závislosti na čase.

Obecně u většiny operátorů bylo potřeba přefiltrovat data v buňkách T_g . Jednak se vyskytla oprava dat typu ex post, dále v některých případech nebyly do výroby vůbec dodány vzorky (to se projevilo jako nulové hodnoty v grafu – vyřešeno jejich eliminací) a stejně tak u některých operátorů se vyskytly dvojí hodnoty teploty skelného přechodu. U operátora PASMA má jedna dávka hmotnost 625 kg. Je evidentní, že tato hodnota dávky vznikla chybou, protože hmotnost ostatních dávek je maximálně do 250 kg. Nelze ani s jistotou říci, jaká hodnota dávky je ta správná, proto dále nebude brána v potaz ve vyhodnocení. Všechny výše popsané závady by se projevily jako chybné hodnoty v popisné statistice.

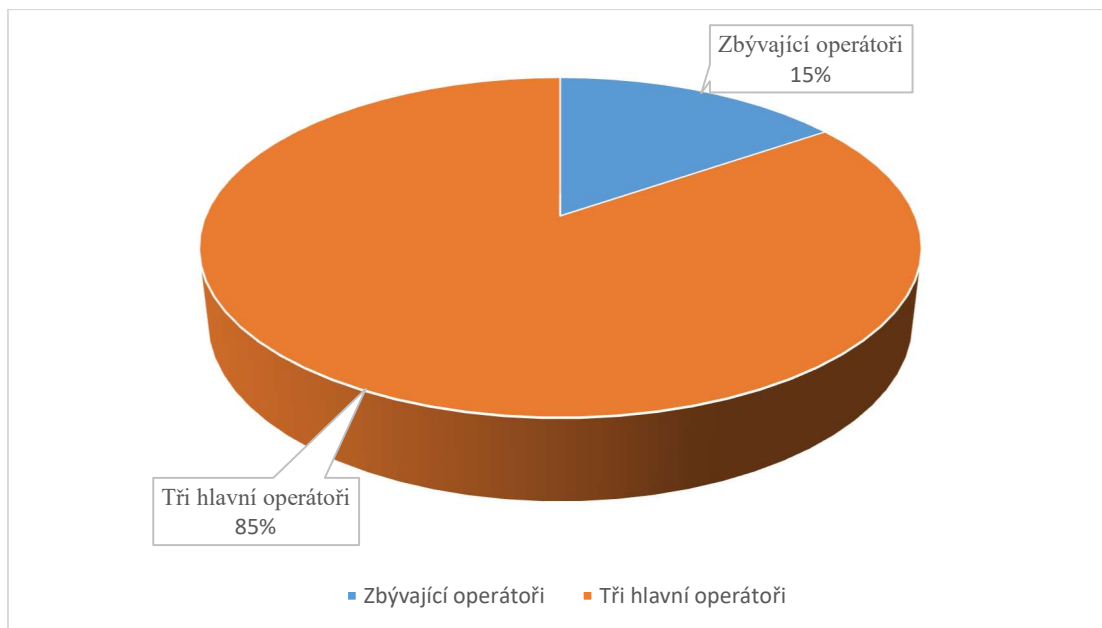
Z tohoto prvotního ohledání vyplynul následující skutkový stav věci:

- existuje skupina tří operátorů,
 - která má větší počet vyrobených kusů oproti ostatním, viz obrázek č. 8;
 - zároveň platí, že hustota rozdělení T_g jejich výrobků je normální;
- existuje skupina pěti operátorů, kteří se na výrobě podílejí pouze sporadicky, a hustota rozdělení T_g jejich výrobků není normální;
- u některých operátorů nabývá hodnota T_g vně minimální výrobní meze;
- u žádného operátora není překročena maximální hodnota výrobní meze T_g ;

- existuje skupina dávek (každá pátá), která koncentruje větší počet vyrobených kusů oproti zbývajícím.

Z tohoto lze vyvodit závěry, že:

- tři operátoři jsou hlavními účastníky výrobního procesu;
- hlavní množství dávky pro výrobu je každá pátá.



Obr. 8 Množství výrobků pro hlavní a vedlejší operátory

Aby mohlo být provedeno náležité posouzení, byly formulovány hlavní statistické hypotézy H_0 , popř. jejich subvarianty a), b):

1. neexistuje statisticky významný vliv faktoru operátora:
 - a) neexistuje statisticky významný rozdíl faktoru tří hlavních operátorů;
 - b) neexistuje statisticky významný rozdíl faktoru zbývajících operátorů;
2. neexistuje statisticky významný rozdíl velikosti dávky.

Dále samozřejmě alternativní hypotézy H_a :

1. existuje statisticky významný vliv faktoru operátora:
 - a) existuje statisticky významný vliv faktoru tří hlavních operátorů;
 - b) existuje statisticky významný vliv faktoru zbývajících operátorů;
2. existuje statisticky významný vliv faktoru dávky.

Pro přehlednost následují tabulky č. 1 a č. 2 s výsledky analýzy ANOVA s hladinou významnosti 0,05 výše uvedených hypotéz:

Tabulka 1 Vliv faktoru dávky

Proměnná	Hodnota p
Vliv faktoru dávky	0,184

Tabulka 2 Vliv faktoru operátor

Proměnná	Hodnota p
Všichni operátoři	0,0001
Tři operátoři	0,2305
Zbývající operátoři	0,0033

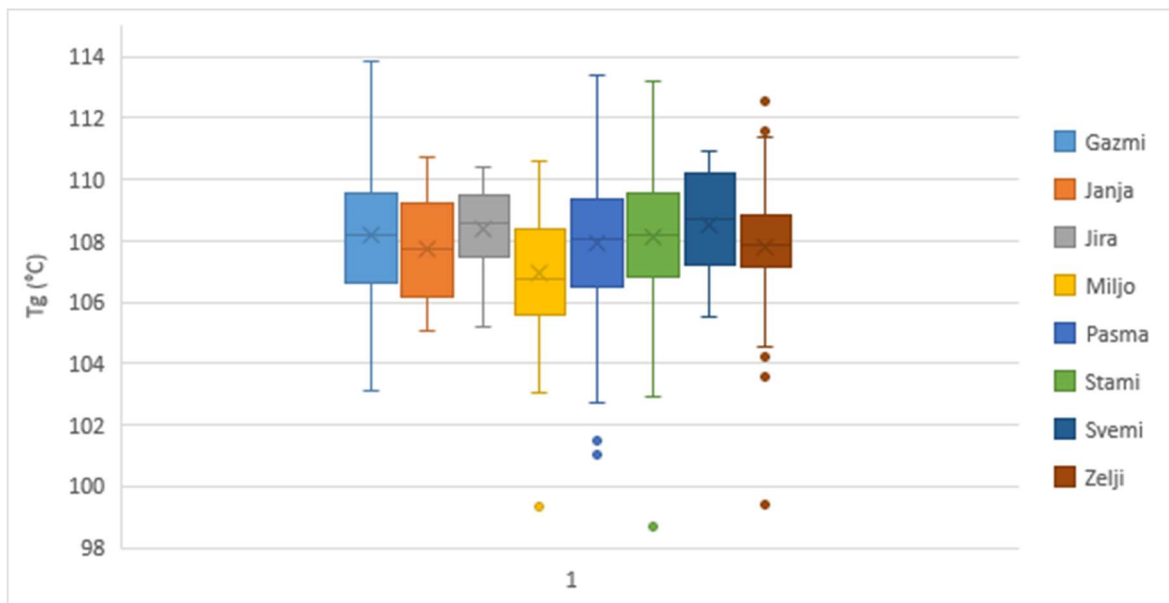
S ohledem na výsledek je vyhodnocení ANOVA analýzy statistických hypotéz bodů 1. a 2. z předchozí stránky provedeno následovně:

- ad 1. zamítáme H_0 a přijímáme H_a ;
- ad 1. a) nezamítáme H_0 . Bod Ad 1) b) zamítáme H_0 a přijímáme H_a ;
- ad 2. nezamítáme H_0 .

V tomto bodě máme přibližnou představu o výrobním procesu. Zjistili jsme, že **existuje** statisticky významný rozdíl mezi všemi a zbývajícími operátory, naopak **neexistuje** statisticky významný rozdíl mezi třemi hlavními operátory.

Pokud jde o hypotézu vlivu velikosti dávky, tak je třeba zdůraznit, že hypotézu hodnotíme ve vztahu k T_g . Ačkoliv H_0 ohledně velikosti dávky nezamítáme, tzn., že neexistuje statisticky významný rozdíl velikosti dávky, tak to stále nepopírá množství kusů v každé páté dávce.

B.1.1 Společné vyhodnocení všech operátorů

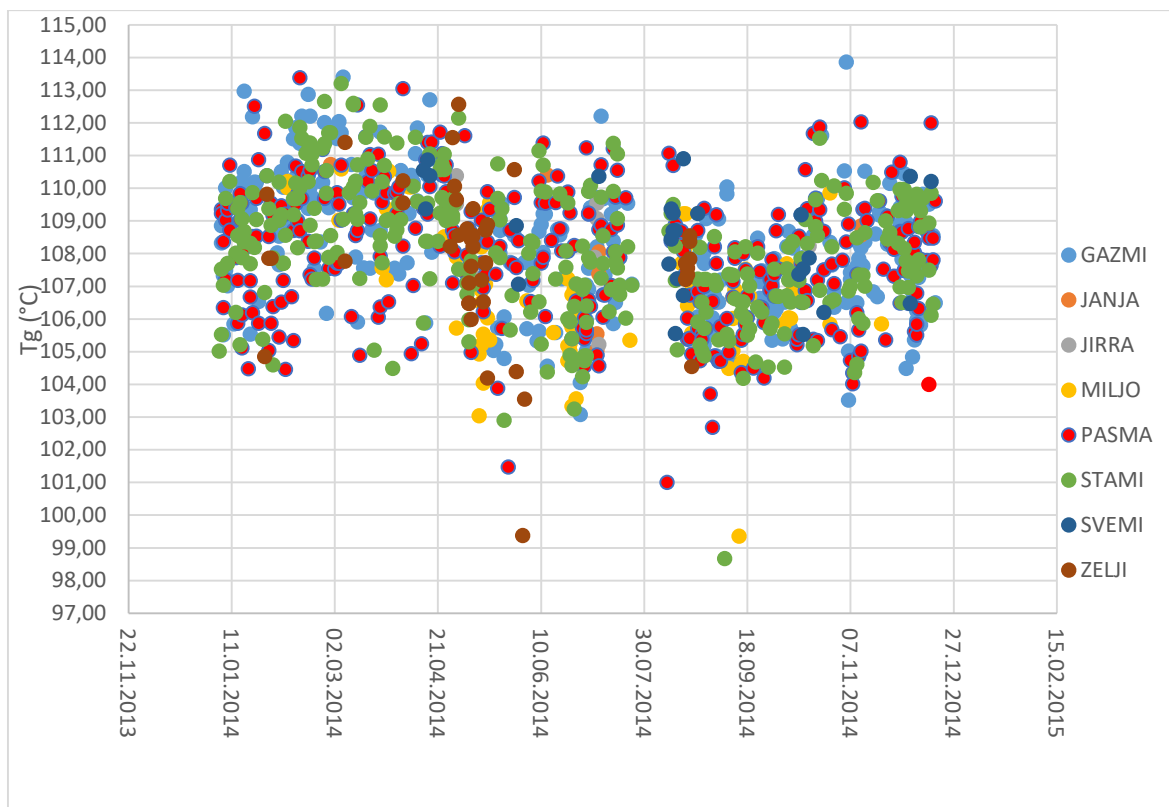


Obr. 9 Krabicový graf všech operátorů

Z krabicového grafu na obrázku č. 9 vplynuly další závěry. Na první pohled je vidět variabilita T_g u hlavních operátorů, která je nejspíše dána celkovým počtem vyrobených kusů. Dále dle mediánové čáry lze pozorovat normální rozdělení pouze u těchto operátorů (bude dále analyzováno). Jak bylo řečeno v úvodu, horní výrobní mez nebyla překročena u žádného operátora. To už nelze prohlásit o mezi dolní, i když je patrné, že poloha dolního kvartilu leží ve všech případech nad výrobní mezí. V tomto bodě lze predikovat, že množství výrobků, které dolní mez překročí, nebude více než 25 %. Zdánlivě se takové číslo může jevit jako vysoké, ale při zavedení výrobku do praxe dochází k jeho dotvrzení a posunu grafu směrem nahoru.

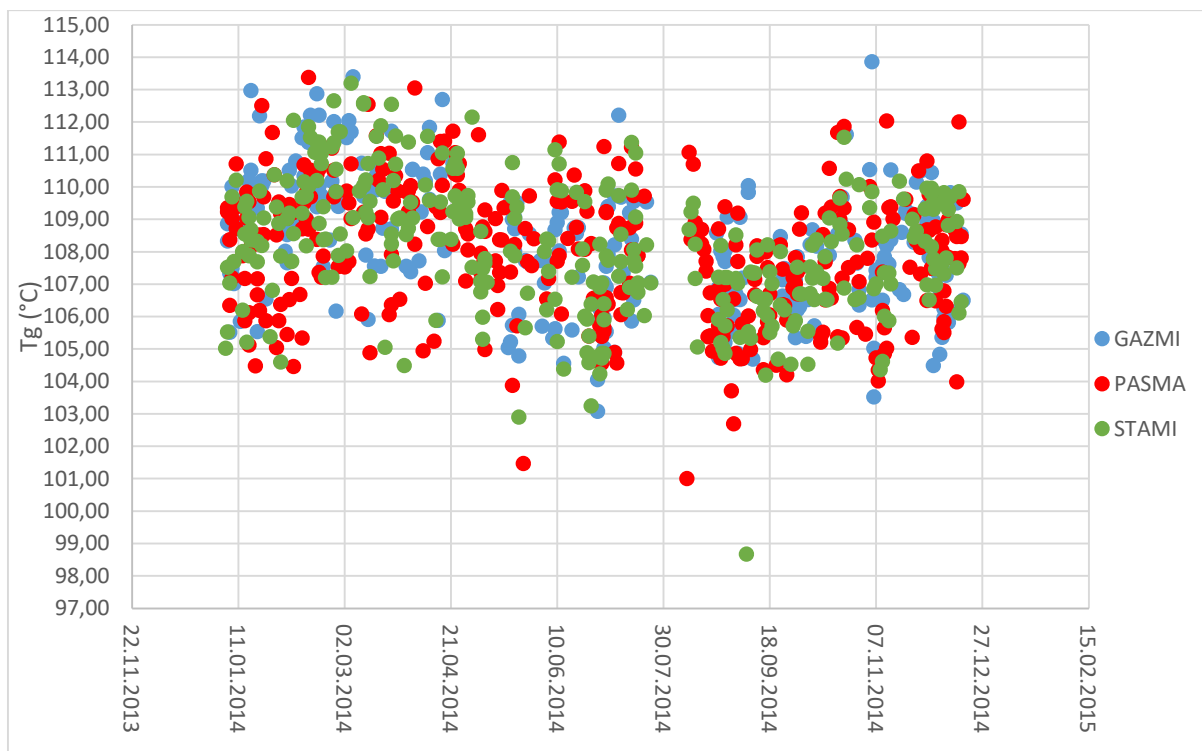
Odlehlé hodnoty obecně nejsou problémem tohoto procesu, když je vidět, že tvoří jen několik hodnot v celém procesu.

Z obrázku č. 10 je evidentní překročení dolní meze u všech hlavních operátorů, nejvíce u operátora PASMA, za ním následují operátoři STAMI a GAZMI.

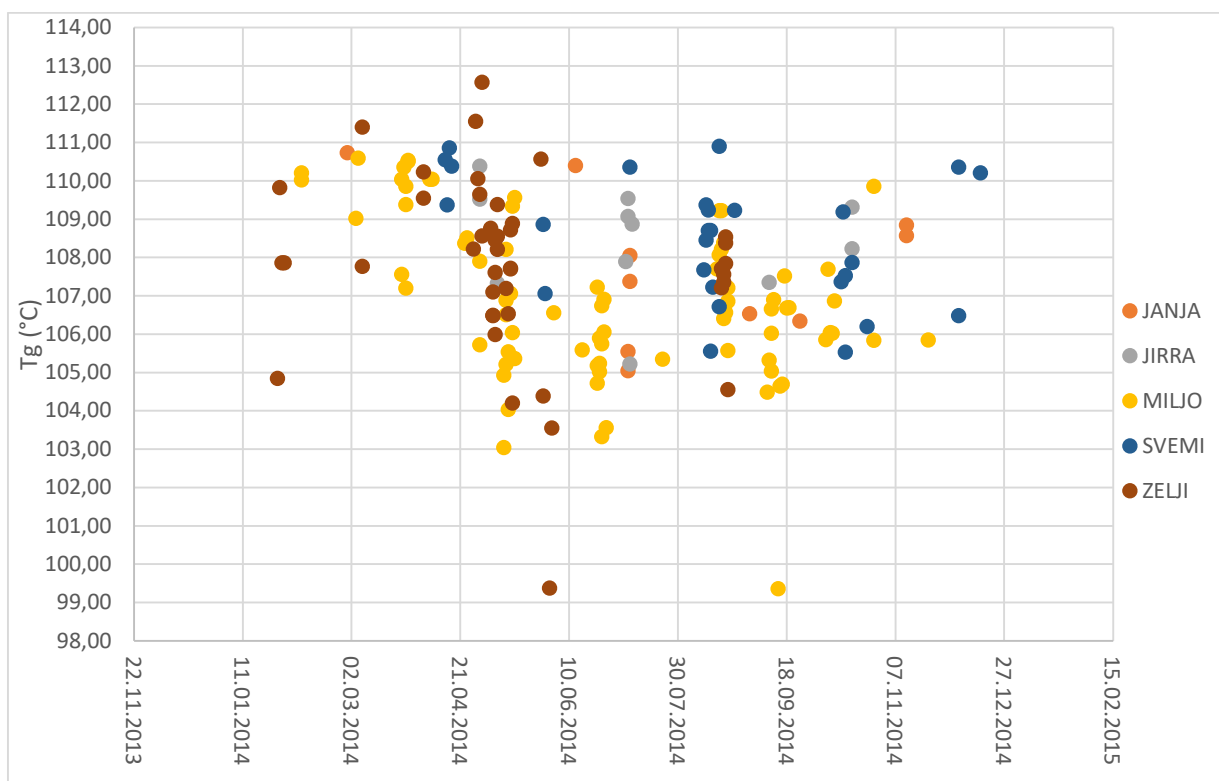


Obr. 10 Hustota množství výrobků pro všechny operátory

Dále je možné pozorovat rozložení hustoty výrobků pro jednotlivé operátory, resp. rozdíl vyrobeného množství u hlavních a zbývajících operátorů. To, jaké míry dosahuje tento rozdíl, znázorňují obrázky č. 11 a 12.



Obr. 11 Hustota množství výrobků pro tři hlavní operátory



Obr. 12 Hustota množství výrobků zbývajících operátorů

Dále provedeme podrobné doanalyzování hlavních operátorů, protože jejich výrobky v součtu tvoří 85 % všech výrobků.

B.1.2 Analýza hlavních operátorů

Tabulka č. 3 zobrazuje vypočítaná základní statistická data u hlavních operátorů. Porovnání hodnot potvrzuje dříve učiněný závěr o neexistenci statisticky významného rozdílu mezi hlavními operátory.

Tabulka 3 Přehled základních statistických hodnot u hlavních operátorů

GAZMI		PASMA		STAMI	
T_g		T_g		T_g	
Stř. hodnota	108,201	Stř. hodnota	107,935	Stř. hodnota	108,129
Chyba stř. hodnoty	0,125	Chyba stř. hodnoty	0,107	Chyba stř. hodnoty	0,113
Medián	108,215	Medián	108,070	Medián	108,195
Modus	108,710	Modus	108,710	Modus	107,040
Směr. odchylka	2,037	Směr. odchylka	2,027	Směr. odchylka	2,025
Rozptyl výběru	4,149	Rozptyl výběru	4,110	Rozptyl výběru	4,100
Špičatost	-0,354	Špičatost	-0,105	Špičatost	0,756
Šikmost	0,128	Šikmost	-0,158	Šikmost	-0,287
Rozdíl max-min	10,780	Rozdíl max-min	12,380	Rozdíl max-min	14,530
Minimum	103,080	Minimum	101,000	Minimum	98,670
Maximum	113,860	Maximum	113,380	Maximum	113,200
Součet	28564,980	Součet	38532,790	Součet	34817,500
Počet	264,000	Počet	357,000	Počet	322,000
Odlehle	13,000	Odlehle	30,000	Odlehle	20,000

Jedná se do určité míry pouze o informativní hodnoty, které nám dají spolu s vyhodnocením dat ze softwaru Statistica odpověď na otázku povahy, resp. regulace procesu. Cílem této podkapitoly je mj. dát odpověď na otázku týkající se povahy procesu, dle které bude vybrán druh regulačního diagramu. Nelze tedy například dále spočítané meze odlehlosti zaměňovat s mezemi regulace (dány výrobními požadavky).

Poté bude následovat grafické zobrazení hustoty výrobků, histogramy a již avizované meze odlehlosti. Vzhledem k podobnému průběhu u všech operátorů bude vyhodnocení provedeno souhrnně pro všechny, s případnými konkrétními poznámkami pro každého operátora.

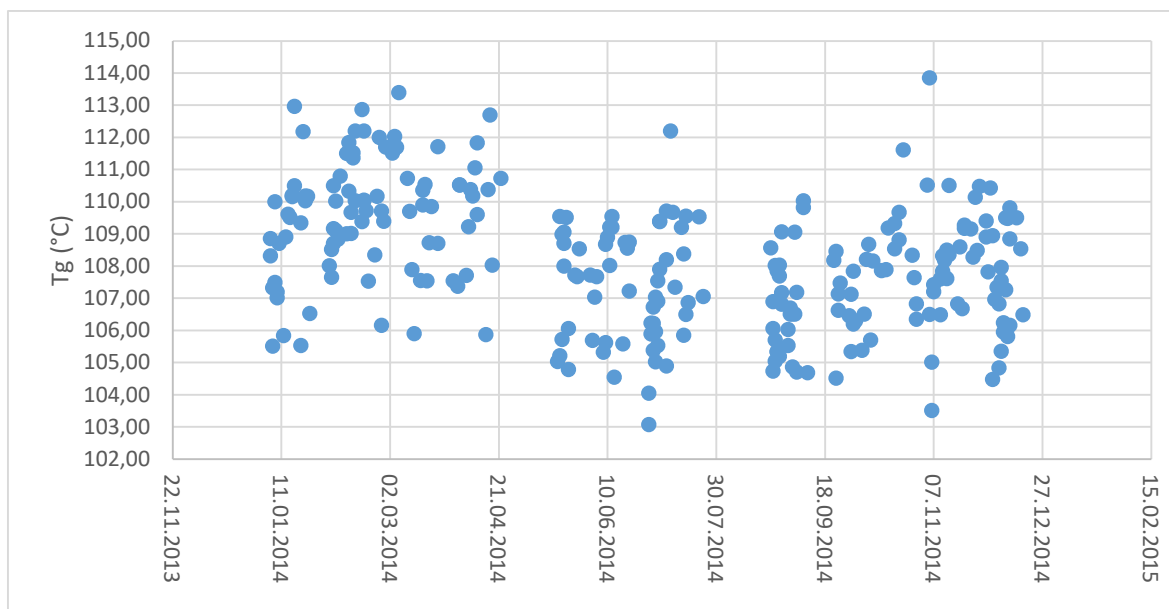
U všech operátorů stojí za povšimnutí trend změny rozložení hustoty výrobků, který bylo možné vidět již na obrázku č. 10. V týdnu přibližně od 21. 04. 2014 dochází k celkovému poklesu maximální T_g výrobků, jež tak setrvává až do konce sledovaného

období. V jednotkách případů dochází po 21. 04. 2014 i k pádu T_g pod mez regulace. Lze vyslovit hypotézu, že došlo k regulaci výroby tak, aby byl využit potenciál materiálu s následným povýrobním vytvrzením, jak bylo popsáno výše.

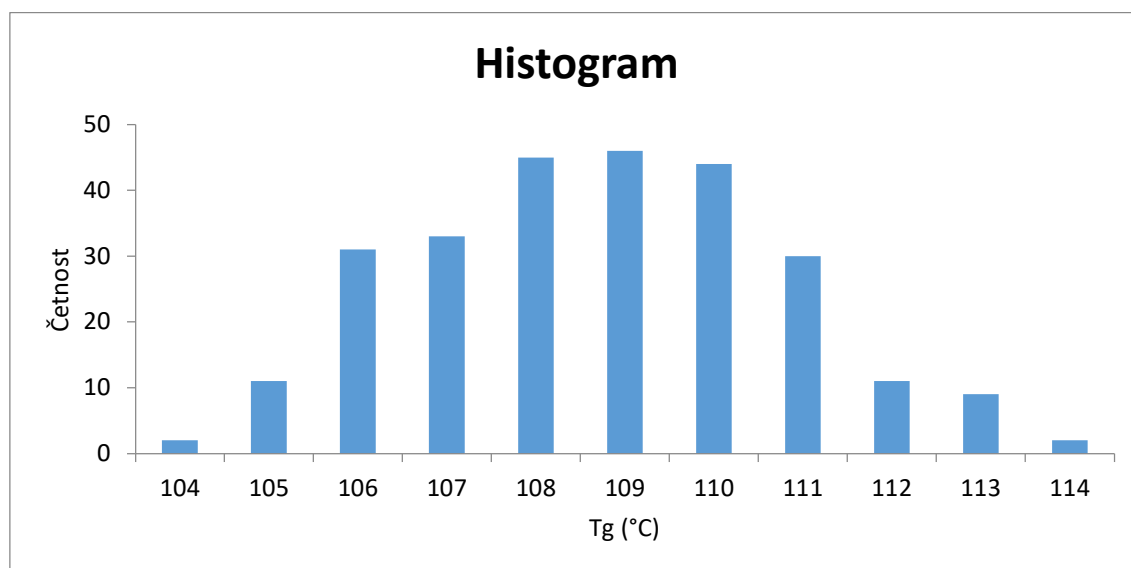
Histogram je u všech operátorů jednovrcholový a jeho průběh značí normální rozdělení. Právě normální rozdělení bude s výhodou využito při další analýze, kde nebude potřeba řešit transformaci dat. Dále můžeme tvrdit, že proces je regulovaný, neboť nedochází k rychlým změnám rozdělení pravděpodobnosti. Dle klasického pojetí SPC by bylo možné vyslovit domněnku, že proces je statisticky zvládnutý. Potřebné vyhodnocení, zda tomu tak skutečně je nebo není spolu s dalšími závěry, bude možné učinit až v následujících kapitolách, zejména po analýze v softwaru Statistica.

B.1.2.1 Operátor GAZMI

Hustota rozložení



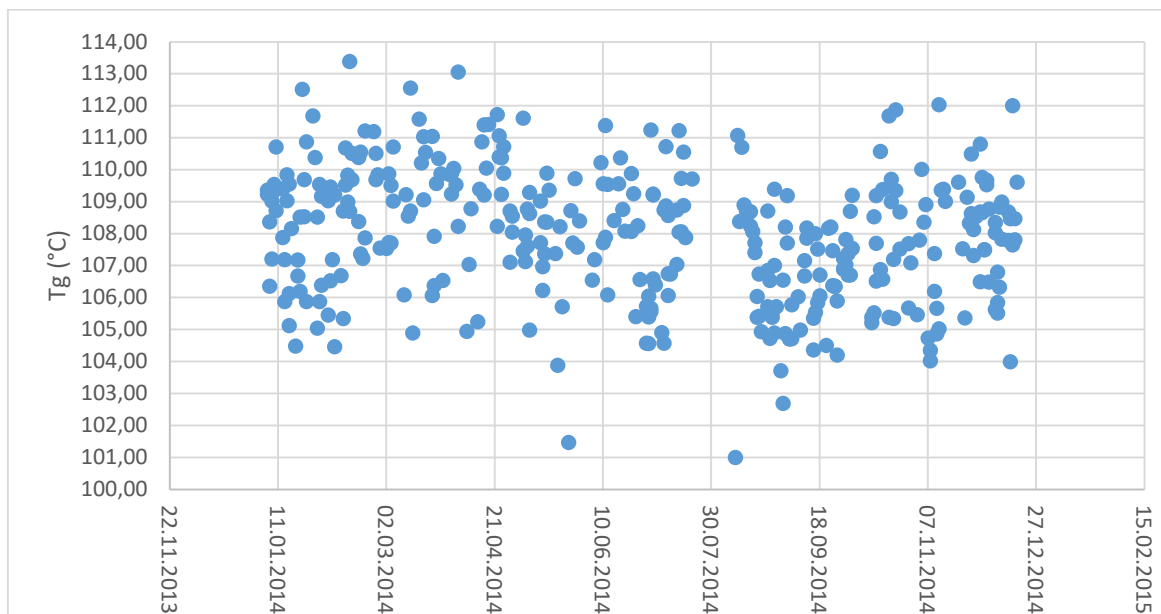
Obr. 13 Hustota rozložení T_g u výrobků u operátora GAZMI

Histogram

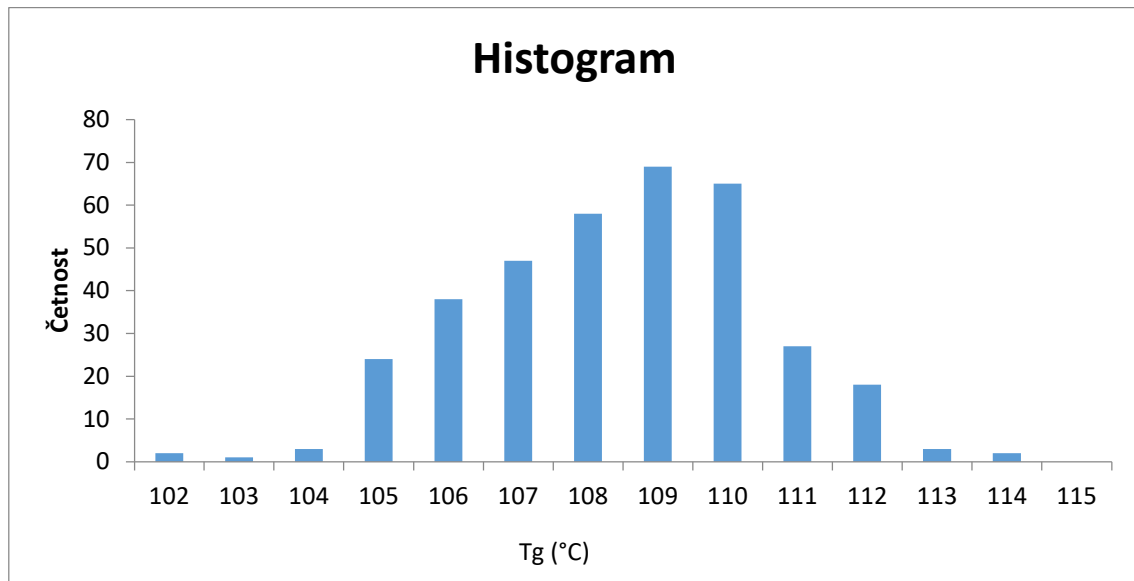
Obr. 14 Histogram Tg výrobků operátora GAZMI

Mez odlehlosti

Metodou $\pm 3\sigma$ vyšla dolní/horní mez cca 102/114 °C.

*B.1.2.2 Operátor PASMA*Hustota rozložení

Obr. 15 Hustota rozložení Tg u výrobků u operátora PASMA

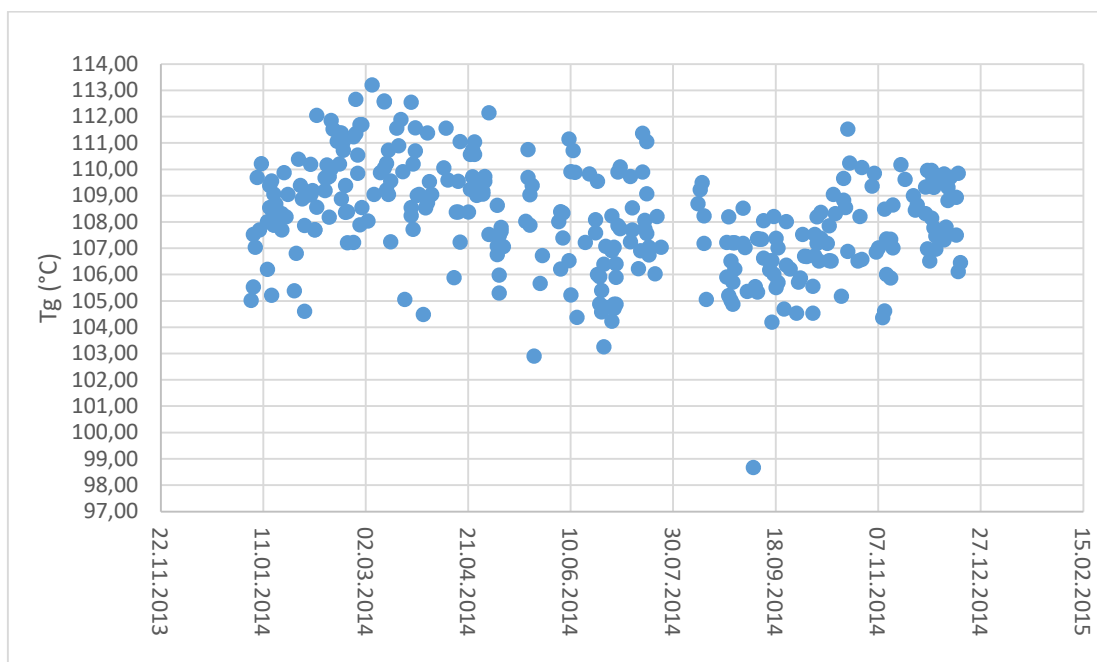
Histogram

Obr. 16 Histogram Tg výrobků operátora PASMA

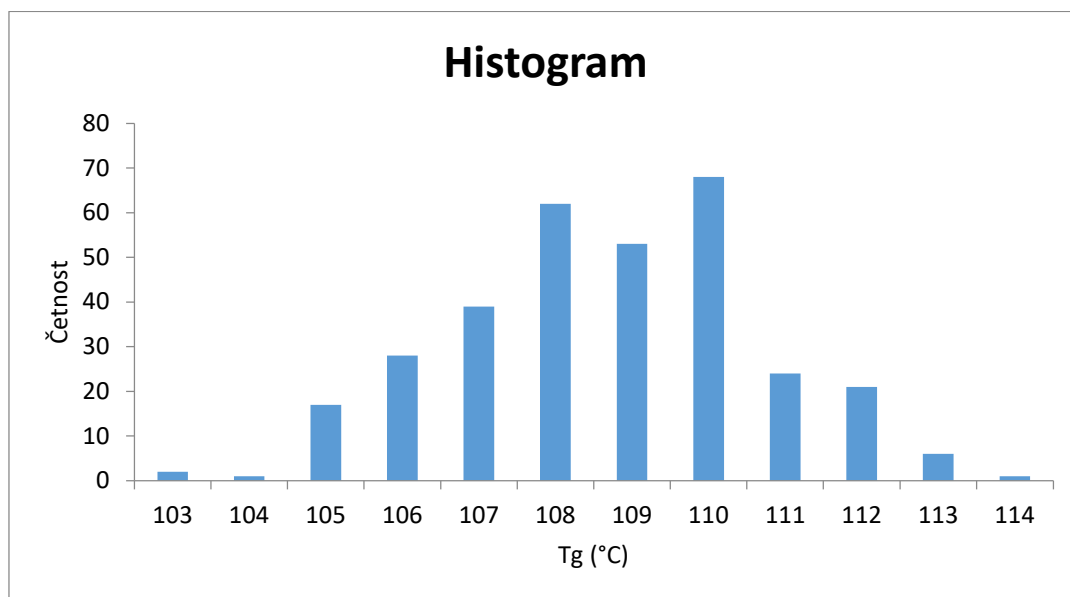
Mez odlehlosti

Metodou $\pm 3\sigma$ vyšla dolní/horní mez cca 102/114 °C.

B.1.2.4 Operátor STAMI

Hustota rozložení

Obr. 17 Hustota rozložení Tg u výrobků u operátora STAMI

Histogram

Obr. 18 Histogram Tg výrobků operátora STAMI

Mez odlehlosti

Metodou $\pm 3\sigma$ vyšla dolní/horní mez cca 102/114 °C.

B.1.3 Analýza zbývajících operátorů

Následující statistická data i pro zbývající operátory rozdělené do dvou tabulek č. 4 a 5. Mezi těmito operátory již vidíme signifikantní diferenci v rozdílu maximální/minimální hodnoty a v rozptylu. Analogicky k vyhodnocení hlavních operátorů. I zde lze vidět oprávněnost závěru o existenci statisticky významného rozdílu mezi zbývajících operátory.

Tabulka 4 Statistická data pro zbývající operátory 1/2-úvod

JANJA		JIRA		MILJO	
<i>T_g</i>		<i>T_g</i>		<i>T_g</i>	
Stř. hodnota	107,745	Stř. hodnota	108,368	Stř. hodnota	106,918
Chyba stř. hodnoty	0,611	Chyba stř. hodnoty	0,400	Chyba stř. hodnoty	0,234
Medián	107,715	Medián	108,550	Medián	106,740
Modus	#NENÍ_K_DISPOZICI	Modus	#NENÍ_K_DISPOZICI	Modus	110,040
Směr. odchylka	1,933	Směr. odchylka	1,384	Směr. odchylka	2,079
Rozptyl výběru	3,735	Rozptyl výběru	1,917	Rozptyl výběru	4,324
Špičatost	-1,006	Špičatost	1,192	Špičatost	1,022
Šikmost	0,240	Šikmost	-0,871	Šikmost	-0,323
Rozdíl max-min	5,680	Rozdíl max-min	5,160	Rozdíl max-min	11,230
Minimum	105,050	Minimum	105,220	Minimum	99,360
Maximum	110,730	Maximum	110,380	Maximum	110,590
Součet	1077,450	Součet	1300,420	Součet	8446,500
Počet	10,000	Počet	12,000	Počet	79,000
Odlehle	0,000	Odlehle	0,000	Odlehle	10

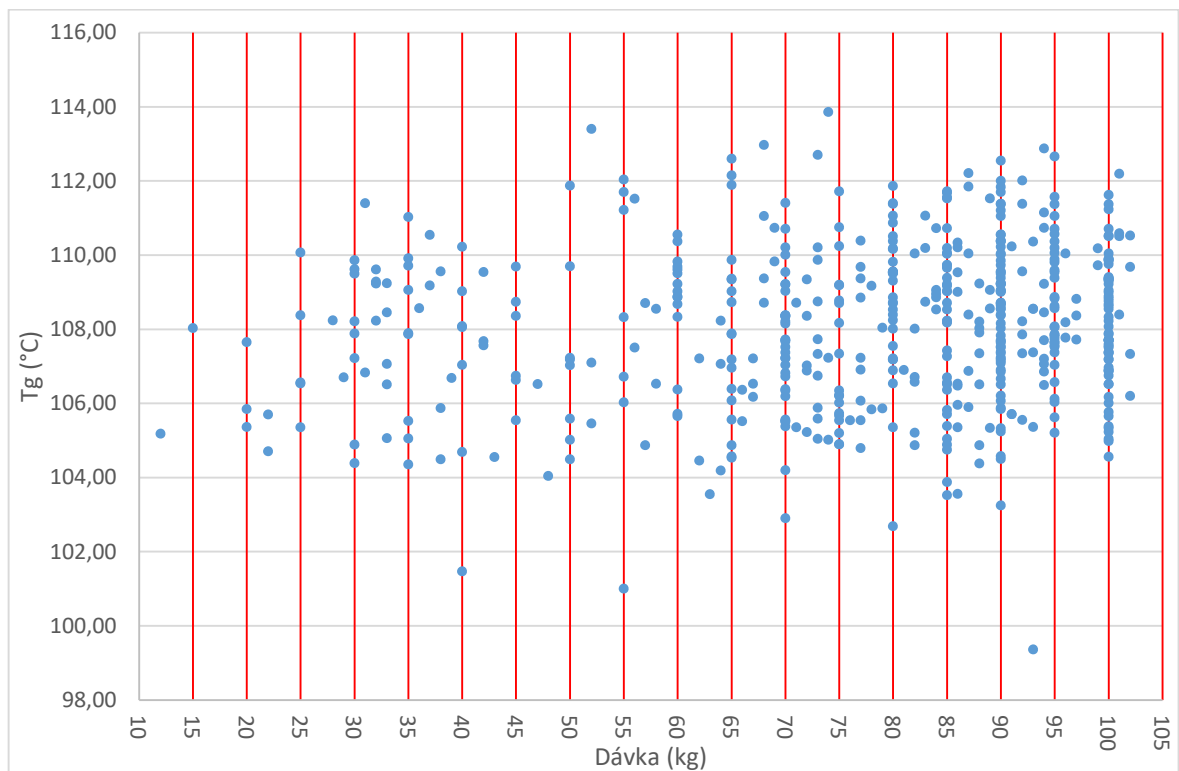
Tabulka 5 Statistická data pro zbývající operátory 2/2-dokončení

SVEMI		ZELJI	
<i>T_g</i>		<i>T_g</i>	
Stř. hodnota	108,517	Stř. hodnota	107,809
Chyba stř. hodnoty	0,314	Chyba stř. hodnoty	0,349
Medián	108,700	Medián	107,860
Modus	109,370	Modus	107,860
Směr. odchylka	1,631	Směr. odchylka	2,316
Rozptyl výběru	2,660	Rozptyl výběru	5,364
Špičatost	-0,998	Špičatost	3,224
Šikmost	-0,256	Šikmost	-1,089
Rozdíl max-min	5,370	Rozdíl max-min	13,190
Minimum	105,530	Minimum	99,380
Maximum	110,900	Maximum	112,570
Součet	2929,960	Součet	4743,590
Počet	27,000	Počet	44,000
Odlehle	0,000	Odlehle	6

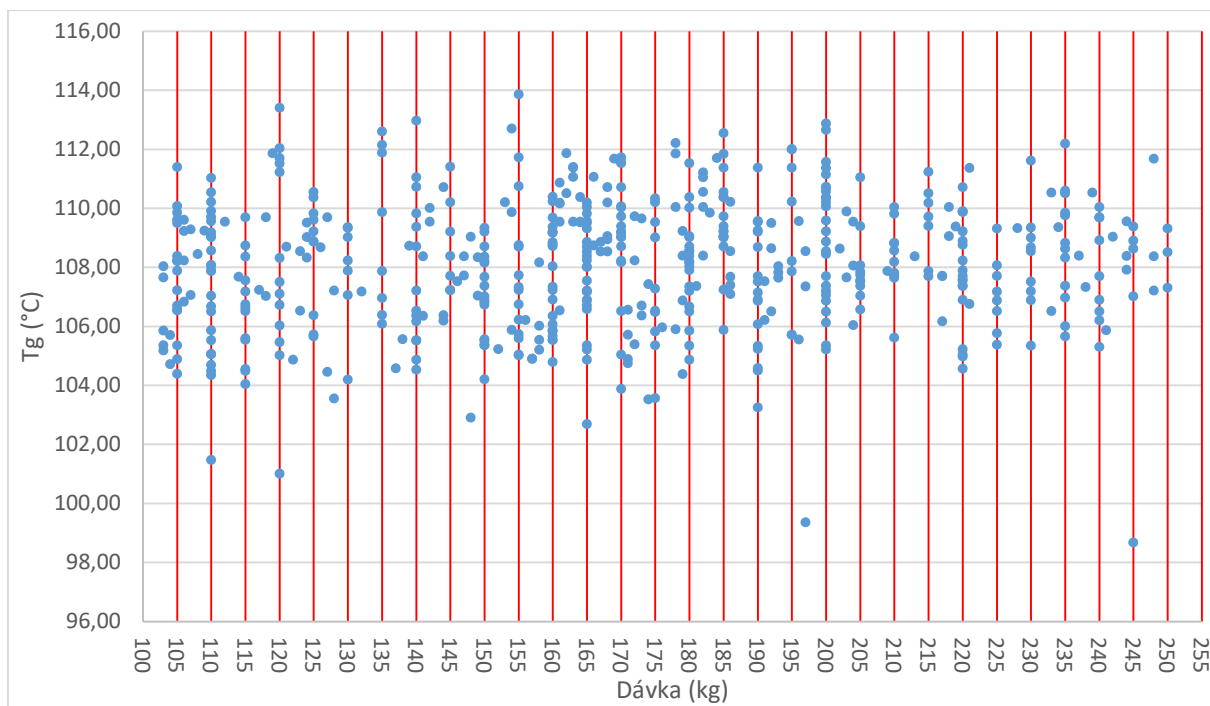
B.1.4 Vyhodnocení vlivu velikosti dávky

Jak bylo uvedeno dříve, každá pátá dávka koncentruje větší množství. Myšlenka je taková, že právě tyto jsou dávkou hlavní výrobní, což potvrzuje grafická interpretace na obrázcích č. 19 a 20.

Pokud jde o vyhodnocení souhrnných statistických dat, tak ty nevykazují významnou vzájemnou odchylku v podskupinách po 10 kg u vyhodnocení T_g . Jinak řečeno, vytvořené podskupiny po 10 kg při porovnání mezi sebou vykazují stejné statistické parametry. Závěr tedy je, že se zde potvrdila statistická hypotéza z úvodu o neexistenci statisticky významného rozdílu velikosti dávky. Vzhledem k rozsahu jsou tabulky č. 6 – 13 s daty v přílohové části E1.



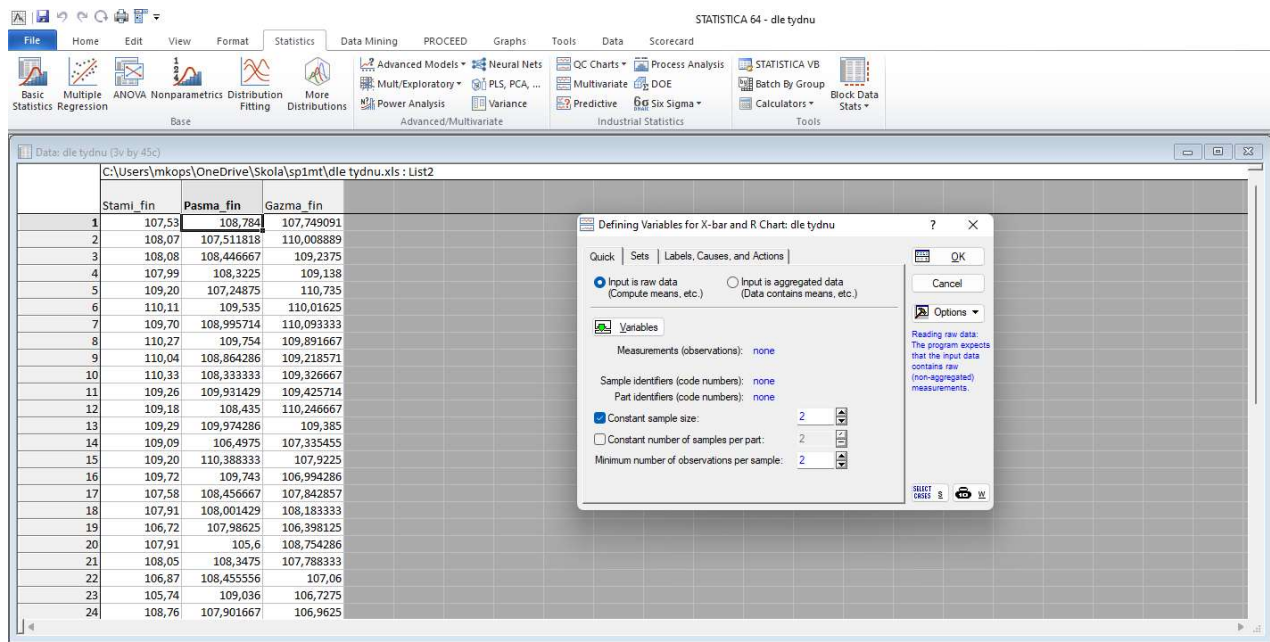
Obr. 19- Vyhodnocení vlivu velikosti dávky 1/2-úvod



Obr. 20 Vyhodnocení vlivu velikosti dávky 2/2-dokončení

B.2 Analýza dat v softwaru Statistica

STATISTICA umožňuje jednoduše importovat listy ze softwaru Excel. Volba proměnných je volitelná pro první hodnotu řádku nebo sloupce. Obrázek č. 21 zachycuje krok v postupu tvorby regulačního diagramu.



Obr. 21 Nabídka tvorby regulačního diagramu

Volba regulačního diagramu

Software nabízí širokou paletu diagramů. S ohledem na teorii diplomové práce byl použit Shewartův regulační diagram pro výběrový průměr a rozptyl. Pro zjednodušení bylo zvoleno hromadné zadání proměnných s nastavením logických podskupin, centrální přímky a regulačních mezí.

Regulační diagramy byly konstruovány pouze pro hlavní operátory vzhledem k normálnímu rozdělení.

Volba znaku jakosti

Volba znaku jakosti vyplynula již ze zadání, je jím teplota skelného přechodu T_g .

Volba logických podskupin

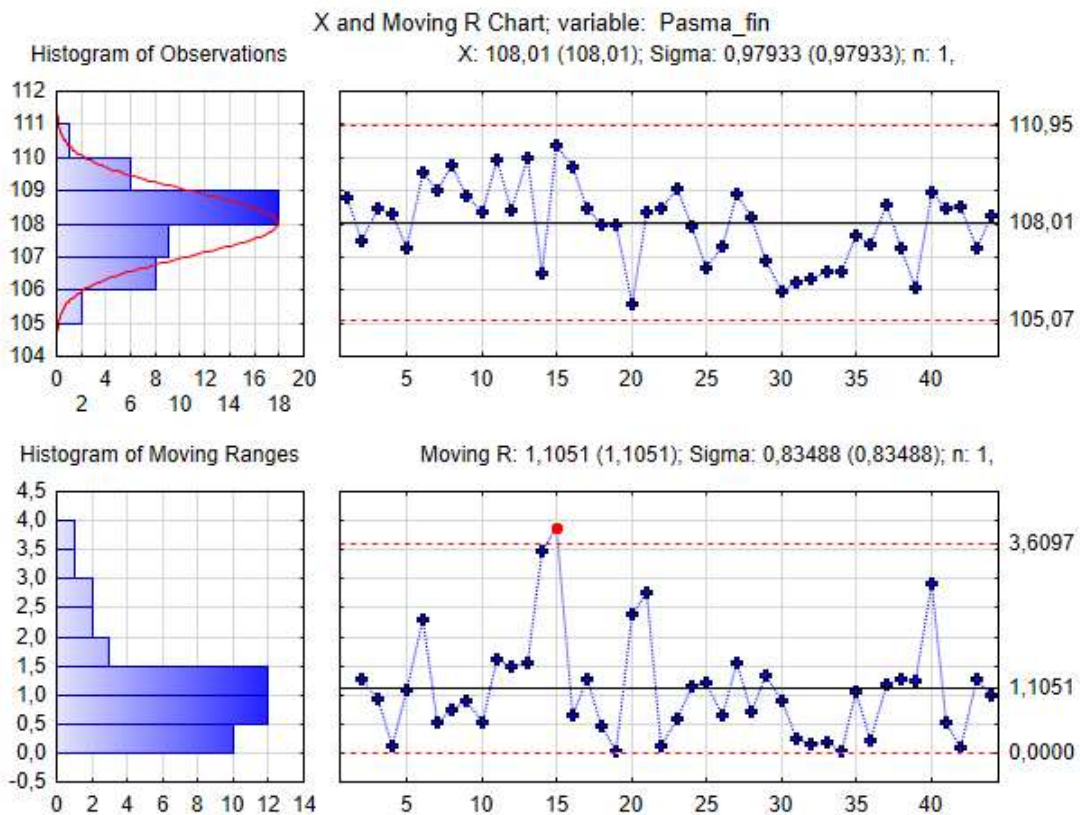
Volba logických podskupin byla provedena po týdnu s ohledem na operátory, resp. na množství jejich výrobků a typ zvoleného diagramu.

Volba centrální přímky a volba meze specifikace

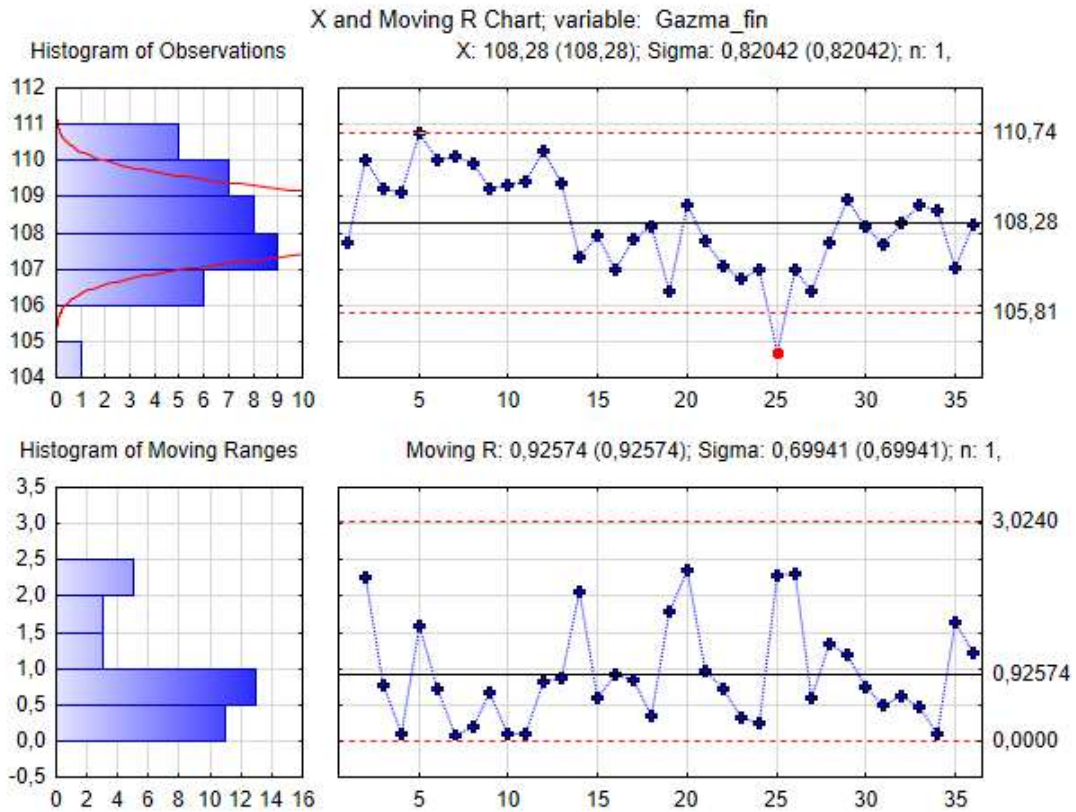
Tyto veličiny jsou vypočítány softwarem STATISTICA. Jejich porovnání s hodnotou stanovenou zadáním (centrální přímka) a hodnotami vypočítanými při práci se softwarem

Excel (regulační meze) dají další odpovědi, nadto by bylo možné spočítat způsobilost výrobního procesu.

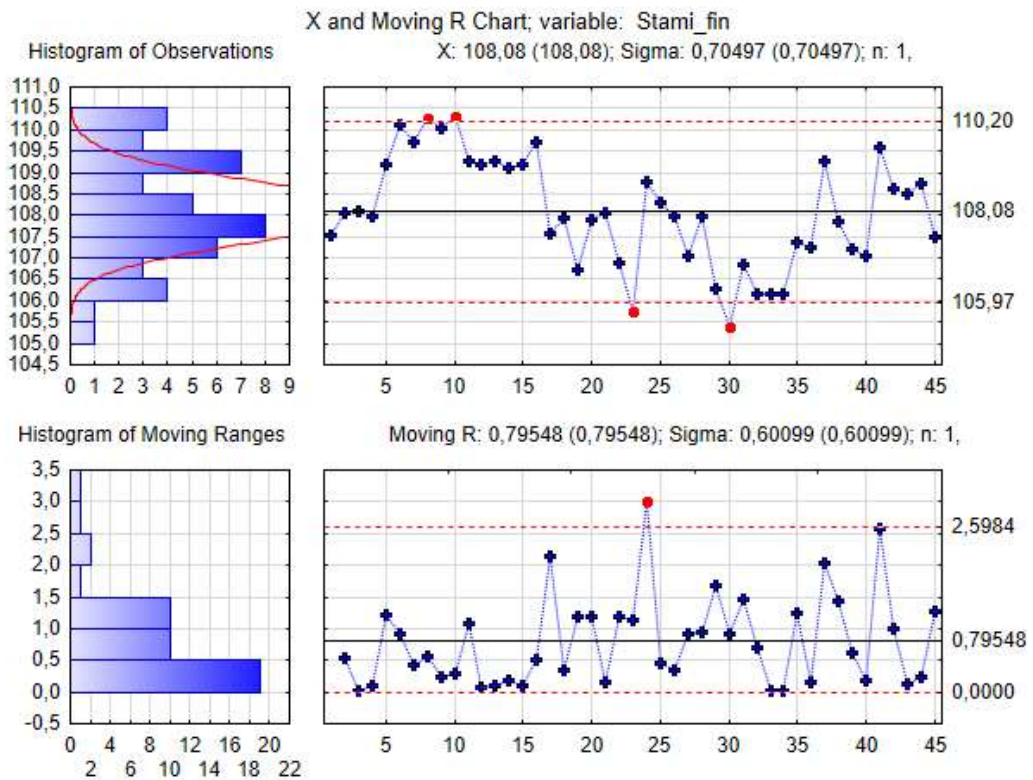
Regulační diagramy jsou na obrázcích č. 22 – 24. Obecně platí, že hodnota centrální přímký je pod požadovanou hodnotou 110 °C. Takový výsledek je správný. Při výrobě se teplota procesu účelně nastaví nižší, neb se předpokládá „dovytvrzení“ materiálu v provozních podmínkách při aplikaci. Důsledkem tohoto „dovytvrzení“ je samozřejmě zvýšení teploty skelného přechodu na požadovaných 110 °C. Výhodou této materiálové charakteristiky je šetření energie při výrobě.



Obr. 22 Regulační diagram operátora PASMA



Obr. 23 Regulační diagram operátora GAZMI



Obr. 24 Regulační diagram operátora STAMI

V kombinaci s detailním vyhodnocením těchto operátorů v předchozím kroku v Excelu lze říci, že proces je nějakým způsobem regulovaný. Usuzovat lze z hodnot a shodnosti všech klíčových parametrů, především pak rozdělení dat, rozptylu, meze odlehlosti spolu s průběhem T_g v jednotlivých týdnech. U všech operátorů je zajímavé pozorovat počáteční polohu T_g nad centrální přímkou. U operátora STAMI dokonce dochází k překročení meze specifikace. Teoreticky by mohlo jít o chybu druhého druhu u regulovaného procesu. Od cca 15. týdne dochází k poklesu a oscilaci okolo centrální přímkou, ačkoliv u operátora STAMI a GAZMI naopak k překmitu na dolní mez. O příčině takového zdánlivě nestálého chování lze na tomto místě pouze spekulovat. Variantou je druh regulace výroby, popřípadě nečistoty v materiálu, které způsobí fluktuaci T_g . Dále stojí za povšimnutí hodnota centrální přímkou 108 °C.

C. Zhodnocení a závěr

Předmětem této diplomové práce byla statistická analýza dat se zaměřením na statistickou regulaci procesu. Práce byla rozdělena na teoretickou a praktickou část. K praktické části byla poskytnuta data z procesu výroby výrobků, ve kterém je použit izolant, který se během výroby vytvrzuje. Z pohledu autora této práce je výroba tvořena celkem osmi operátory a výrobky, které se procesem vytvrzují na požadovanou teplotu skelného přechodu T_g 110 °C, přičemž dovolená výrobní tolerance je v mezích ± 5 °C.

Prvním nezbytným krokem bylo rozřídění, resp. filtrace poskytnutých dat. Filtrací byla odstraněna extrémní data a jejich rozříděním došlo k elementárnímu vyhodnocení výroby. Tím bylo především zjištěno, že existuje skupina tří operátorů, která má větší počet vyrobených kusů oproti ostatním, a hustota rozdělení T_g jejich výrobků je normální. Na druhé straně zbývající skupina pěti operátorů se na výrobě podílí pouze sporadicky, přičemž hustota rozdělení T_g jejich výrobků normální není. Nepřekročení horní výrobní meze je společným faktorem všech operátorů, u některých dochází k překročení meze dolní. Dalším podstatným závěrem je, že každá pátá dávka koncentruje větší množství výrobků.

K ověření všech těchto tvrzení byly formulovány jednak statistické hypotézy s hladinou významnosti 0,05 (viz tabulky č. 1 a č. 2 vyhodnocení faktoru dávky/operátora) a jednak byla provedena společná analýza operátorů v Excelu (viz část B 1.1.). Z části B 1.1. stojí za zmínku poloha dolního kvartilu. Ta ve všech případech leží nad výrobní mezí. Lze tedy vzít v úvahu, že množství výrobků, které dolní mez překročí, nebude větší než 25 %. Zdánlivě se takové číslo může jevit jako vysoké, ale při zavedení výrobku do praxe dochází k jeho dotvrzení a posunu grafu směrem nahoru.

Pro ověření dříve stanovených předpokladů a stanovení postupu v navazující části analýzy, tedy například pro výběr vhodného regulačního diagramu nebo určení toho, zda je nebo není nutné data transformovat, byly provedeny podrobnější analýzy. Analýza byla provedena pro jednotlivé skupiny operátorů dohromady, v případě hlavních operátorů pak u každého samostatně. Jak bylo uvedeno výše, kritérium pro analýzu je T_g , ale i množství dávky (v kg) dodané k vytvrzení.

V závěrečné etapě diplomové práce bylo přistoupeno k vyhodnocení hlavních operátorů v softwaru Statistica. Volba znaku jakosti, centrální přímky a meze specifikace vplynula již ze zadání. Jako regulační diagram byl zvolen Shewartův regulační diagram

pro výběrový průměr a rozptyl. Z analýzy grafů v návaznosti na předchozí vyhodnocení v softwaru Excel (při porovnání například shodnosti důležitých parametrů jako jsou rozdělení dat, rozptylu nebo meze odlehlosti spolu s průběhem T_g v jednotlivých týdnech) lze říci, že dochází k regulaci tohoto výrobního procesu. U všech operátorů pozorujeme počáteční polohu T_g nad centrální přímkou. U operátora STAMI navíc dochází k překročení výrobní meze. Jestli se jedná např. o chybu 2. druhu u regulovaného procesu, nelze bez dalších informací o procesu určit. Dále od 15. týdne dochází u všech operátorů k poklesu a oscilaci okolo centrální přímky, ačkoli u operátorů STAMI a GAZMI naopak k překmitu nad dolní mez. Příčinami takového nestálého chování by mohl být druh regulace výroby, popřípadě nečistoty v materiálu, které způsobí fluktuaci T_g . Hodnota centrální přímky (CL) 108 °C u všech operátorů, také svědčí pro dříve učiněný závěr o povýrobním vytvrzení materiálu.

Dle regulačních diagramů lze výrobní proces celkově zhodnotit jako statisticky zvládnutý. Nicméně je zde prostor pro taková opatření, aby ani v jednotkách případů nedocházelo k překročení výrobní meze. Například užší volba intervalu logických podskupin tak, aby docházelo k zachycení nežádaných „posuvů“ v rámci procesu a to v reálném čase.

D. Literatura

- [1] M. Hammer a J. Champy, *Reengineering - radikální proměna firmy: manifest revoluce v podnikání*, 3. vyd. Praha: Management Press, 2000.
- [2] M. Botek a L. Adamec, *Sbírka příkladů z inženýrské ekonomiky a managementu*, 1. vyd. Praha: VŠCHT Praha, 2004.
- [3] E. Cs. Doc. Ing. Jarošová a D. Cs. prof. Ing. Noskovičová, *Pokročilejší metody statistické regulace procesu*. Praha: Grada Publishing, a.s., 2015.
- [4] J. Hendl, *Přehled statistických metod zpracování dat: Analýza a metaanalýza dat*, 4. vyd. Praha: Portál, 2012.
- [5] J. Tošenovský a D. Noskovičová, *Statistické metody pro zlepšování jakosti*. Ostrava: Montanex, 2000.
- [6] J. CSc. Doc. Ing. Kožíšek a B. Ph. D. Ing. Stieberová, *Statistická a rozhodovací analýza*. Praha: České vysoké učení technické v Praze, 2014.
- [7] S. Richterová, „Využití statistických metod ve výrobním procesu společnosti Tegü VUKO, spol. s r.o.“, Bakalářská práce, Univerzita Tomáše Bati ve Zlíně, Zlín, 2010. Viděno: 12. únor 2023. [Online]. Dostupné z: https://digilib.k.utb.cz/bitstream/handle/10563/14105/richterov%3%a1_2010_bp.pdf?sequence=1&isAllowed=y
- [8] M. Bc. Peroutka, „Využití statistických metod pro zajištění způsobilosti procesu výroby“, Diplomová práce, Vysoké učení technické v Brně, Brno, 2016. Viděno: 12. únor 2023. [Online]. Dostupné z: <https://dspace.vutbr.cz/handle/11012/66337>
- [9] P. Bc. Klapka, „Využití nástrojů kvality v procesu uvolnění nové výrobní linky při montáži bezpečnostních pásů“, Diplomová práce, Technická univerzita v Liberci, Liberec, 2014. Viděno: 12. únor 2023. [Online]. Dostupné z: https://dspace.tul.cz/bitstream/handle/15240/124251/V_18914_T.pdf?sequence=-1
- [10] J. Ph. D. Ing. Zmatlík a P. Ing. Zdvořák, „Problémy hodnocení výkonnosti a způsobilosti řízení procesů v rámci nesplnění normality rozdělení dominantního znaku jakosti“, Praha, srp. 2018. Viděno: 12. únor 2023. [Online]. Dostupné z: <https://docplayer.cz/151584958-Problemy-hodnoceni-vykonnosti-a-zpusobilosti-rizeni-procesu-v-ramci-nesplneni-normality-rozdeleni-dominantniho-znaku-jakosti.html>

- [11] M. Kuchovský, „Analýza výrobního procesu v podniku“, Diplomová práce, Masarykova univerzita, Brno, 2016. Viděno: 12. únor 2023. [Online]. Dostupné z: https://is.muni.cz/th/kw1tu/DP_kuchovsky_analyza_vyrobniho_procesu_v_podniku_final_verze.pdf
- [12] M. Hlavatá, „Použití matematicko-statistických metod při řízení procesů ve výrobním podniku HMMC, s.r.o“, Bakalářská práce, Univerzita Tomáše Bati ve Zlíně, Zlín, 2012. Viděno: 12. únor 2023. [Online]. Dostupné z: http://digilib.k.utb.cz/bitstream/handle/10563/22441/hlavat%c3%a1_2012_bp.pdf?sequence=1&isAllowed=y
- [13] P. Bc. Horník, „Matematické modely způsobilosti procesu“, Diplomová práce, Vysoké učení technické v Brně, Brno, 2015. Viděno: 12. únor 2023. [Online]. Dostupné z: https://www.vut.cz/www_base/zav_prace_soubor_verejne.php?file_id=103349
- [14] V. Vassilenko, „Statistická analýza a regulace výrobního procesu v podniku“, Diplomová práce, České vysoké učení technické, Praha, 2018. Viděno: 12. únor 2023. [Online]. Dostupné z: https://dspace.cvut.cz/bitstream/handle/10467/80576/F2-DP-2018-Vassilenko-Viktor-DP_Viktor_Vassilenko.pdf?sequence=-1&isAllowed=y
- [15] K. Kupka, *Statistické řízení jakosti*. Pardubice: TriloByte, 1997.
- [16] O. Tůmová a D. Pirich, *Nástroje řízení jakosti a základy technické diagnostiky*. Plzeň: Západočeská univerzita v Plzni, 2003.
- [17] P. Tvrdý, „Regulační diagramy řešené SW QI Analyst“, Diplomová práce, Západočeská univerzita v Plzni, Plzeň, 2013. Viděno: 12. únor 2023. [Online]. Dostupné z: https://dspace5.zcu.cz/bitstream/11025/7480/1/TVRDY_E11N0042P_DP.pdf
- [18] M. Krpenská, „Využití regulačních diagramů v monitoringu kvality zdravotního screeningu“, Bakalářská práce, Masarykova univerzita, Brno, 2014. Viděno: 12. únor 2023. [Online]. Dostupné z: https://is.muni.cz/th/qn0vh/BP_Michaela_Krpenska.pdf
- [19] L. Ph. D. Ing. Kupka a V. Ing. Kubernát, „Aplikace příkladů z normy ČSN ISO 8258 Shewhartovy regulační diagramy do programu QC.Expert 2.5“, *ElectroScope*, Plzeň, s. 1–4, 2008. Viděno: 12. únor 2023. [Online]. Dostupné z: http://147.228.94.30/images/PDF/Rocnik2008/cislo2_2008/r0c1c4.pdf

- [20] D. Montgomery, *Introduction to Statistical quality control*, 6. vyd. Elm Street Publishing Services, 2009. Viděno: 14. květen 2023. [Online]. Dostupné z: <https://www.uaar.edu.pk/fs/books/12.pdf>
- [21] L. Kotlorz, „Testy normality“, Bakalářská práce, Univerzita Karlova v Praze, Praha, 2012. Viděno: 12. únor 2023. [Online]. Dostupné z: https://dspace.cuni.cz/bitstream/handle/20.500.11956/45596/BPTX_2011_1__0_284_184_0_114348.pdf?sequence=1&isAllowed=y
- [22] E. C. Doc. Ing. Jarošová a J. Cs. Ing. Král, „Ověřování předpokladu normality“, 2010. Viděno: 12. únor 2023. [Online]. Dostupné z: <https://adoc.pub/narodniinformanistedisko-pro-podporu-jakosti.html>
- [23] M. Haičman, „Simulace a testy normality“, Bakalářská práce, Masarykova univerzita, Brno, 2011. Viděno: 12. únor 2023. [Online]. Dostupné z: <https://is.muni.cz/th/yncl0/simultestynormality.pdf>
- [24] kolektiv autorů, „Matlab Instalační manuál“, 2022. Viděno: 13. květen 2023. [Online]. Dostupné z: https://www.mathworks.com/help/matlab/index.html?s_tid=hc_panel
- [25] kolektiv autorů, „SAS Instalační manuál“, 2022. Viděno: 13. květen 2023. [Online]. Dostupné z: http://documentation.sas.com/doc/en/pgmsascdc/9.4_3.4/statug/titlepage.htm
- [26] kolektiv autorů, „STATA Instalační manuál“, 2022. Viděno: 13. květen 2023. [Online]. Dostupné z: <https://www.stata.com/manuals/u.pdf>
- [27] kolektiv autorů, „IBM SPSS Instalační manuál“, 2022. Viděno: 13. květen 2023. [Online]. Dostupné z: https://www.ibm.com/docs/en/SSLVMB_29.0.0/pdf/GPL_Reference_Guide_for_IBM_SPSS_Statistics.pdf
- [28] kolektiv autorů, „TIBCO Statistica Instalační manuál“, 2022. Viděno: 13. květen 2023. [Online]. Dostupné z: https://docs.tibco.com/pub/stat/14.0.1/doc/pdf/TIB_stat_14.0.1_installation.pdf?id=0

E. Přílohy

E.1 Vyhodnocení vlivu velikosti dávky

Tabulka 6 Statistická data dle jednotlivých velikostí dávek 1/8 - úvod

<i>12-20</i>		<i>22-30</i>		<i>31-40</i>	
Stř. hodnota	106,414	Stř. hodnota	107,284	Stř. hodnota	107,787
Chyba stř. hodnoty	0,595	Chyba stř. hodnoty	0,456	Chyba stř. hodnoty	0,384
Medián	105,850	Medián	107,220	Medián	108,155
Modus	není k dispozici	Modus	není k dispozici	Modus	není k dispozici
Směr. odchylka	1,331	Směr. odchylka	1,879	Směr. odchylka	2,239
Rozptyl výběru	1,773	Rozptyl výběru	3,531	Rozptyl výběru	5,012
Špičatost	-2,906	Špičatost	-1,253	Špičatost	0,397
Šikmost	0,508	Šikmost	-0,017	Šikmost	-0,749
Rozdíl max-min	2,850	Rozdíl max-min	5,680	Rozdíl max-min	9,930
Minimum	105,180	Minimum	104,390	Minimum	101,470
Maximum	108,030	Maximum	110,070	Maximum	111,400
Součet	532,070	Součet	1823,820	Součet	3664,770
Počet	5,000	Počet	17,000	Počet	34,000

Tabulka 7 Statistická data dle jednotlivých velikostí dávek 2/8 - pokračování

<i>42-50</i>		<i>52-60</i>		<i>62-70</i>	
Stř. hodnota	107,185	Stř. hodnota	108,418	Stř. hodnota	107,676
Chyba stř. hodnoty	0,457	Chyba stř. hodnoty	0,455	Chyba stř. hodnoty	0,276
Medián	107,105	Medián	108,700	Medián	107,445
Modus	není k dispozici	Modus	není k dispozici	Modus	107,210
Směr. odchylka	2,043	Směr. odchylka	2,533	Směr. odchylka	2,243
Rozptyl výběru	4,173	Rozptyl výběru	6,418	Rozptyl výběru	5,030
Špičatost	-0,077	Špičatost	1,176	Špičatost	-0,222
Šikmost	0,445	Šikmost	-0,596	Šikmost	0,276
Rozdíl max-min	7,830	Rozdíl max-min	12,400	Rozdíl max-min	10,070
Minimum	104,040	Minimum	101,000	Minimum	102,900
Maximum	111,870	Maximum	113,400	Maximum	112,970
Součet	2143,700	Součet	3360,970	Součet	7106,610
Počet	20,000	Počet	31,000	Počet	66,000

Tabulka 8 Statistická data dle jednotlivých velikostí dávek 3/8 - pokračování

<i>71-80</i>		<i>81-90</i>		<i>91-100</i>	
Stř. hodnota	108,074	Stř. hodnota	108,229	Stř. hodnota	108,322
Chyba stř. hodnoty	0,241	Chyba stř. hodnoty	0,180	Chyba stř. hodnoty	0,177
Medián	108,250	Medián	108,530	Medián	108,220
Modus	105,540	Modus	110,040	Modus	107,370
Směr. odchylka	2,168	Směr. odchylka	2,114	Směr. odchylka	1,973
Rozptyl výběru	4,702	Rozptyl výběru	4,468	Rozptyl výběru	3,893
Špičatost	-0,383	Špičatost	-0,568	Špičatost	2,394
Šikmost	0,143	Šikmost	-0,204	Šikmost	-0,604
Rozdíl max-min	11,170	Rozdíl max-min	9,300	Rozdíl max-min	13,510
Minimum	102,690	Minimum	103,250	Minimum	99,360
Maximum	113,860	Maximum	112,550	Maximum	112,870
Součet	8753,960	Součet	14935,630	Součet	13431,940
Počet	81,000	Počet	138,000	Počet	124,000

Tabulka 9 Statistická data dle jednotlivých velikostí dávek 4/8 - pokračování

<i>101-110</i>		<i>111-120</i>		<i>121-130</i>	
Stř. hodnota	107,985	Stř. hodnota	107,976	Stř. hodnota	107,763
Chyba stř. hodnoty	0,284	Chyba stř. hodnoty	0,355	Chyba stř. hodnoty	0,316
Medián	108,105	Medián	108,210	Medián	107,730
Modus	109,550	Modus	není k dispozici	Modus	není k dispozici
Směr. odchylka	2,310	Směr. odchylka	2,037	Směr. odchylka	1,702
Rozptyl výběru	5,334	Rozptyl výběru	4,151	Rozptyl výběru	2,896
Špičatost	2,828	Špičatost	-0,160	Špičatost	-0,215
Šikmost	-0,814	Šikmost	0,270	Šikmost	0,214
Rozdíl max-min	14,380	Rozdíl max-min	8,540	Rozdíl max-min	7,050
Minimum	98,670	Minimum	104,020	Minimum	104,530
Maximum	113,050	Maximum	112,560	Maximum	111,580
Součet	7127,010	Součet	3563,220	Součet	3125,120
Počet	66,000	Počet	33,000	Počet	29,000

Tabulka 10 Statistická data dle jednotlivých velikostí dávek 5/8 - pokračování

<i>132-140</i>		<i>141-150</i>		<i>151-160</i>	
Stř. hodnota	107,970	Stř. hodnota	107,554	Stř. hodnota	107,769
Chyba stř. hodnoty	0,395	Chyba stř. hodnoty	0,389	Chyba stř. hodnoty	0,262
Medián	108,180	Medián	107,520	Medián	107,740
Modus	není k dispozici	Modus	není k dispozici	Modus	108,720
Směr. odchylka	2,055	Směr. odchylka	2,365	Směr. odchylka	1,833
Rozptyl výběru	4,222	Rozptyl výběru	5,593	Rozptyl výběru	3,360
Špičatost	0,056	Špičatost	2,897	Špičatost	-0,026
Šikmost	-0,621	Šikmost	-0,772	Šikmost	0,251
Rozdíl max-min	7,990	Rozdíl max-min	12,670	Rozdíl max-min	8,350
Minimum	103,080	Minimum	99,380	Minimum	104,200
Maximum	111,070	Maximum	112,050	Maximum	112,550
Součet	2915,200	Součet	3979,510	Součet	5280,670
Počet	27,000	Počet	37,000	Počet	49,000

Tabulka 11 Statistická data dle jednotlivých velikostí dávek 6/8 - pokračování

<i>161-170</i>		<i>171-180</i>		<i>181-190</i>	
Stř. hodnota	107,926	Stř. hodnota	108,279	Stř. hodnota	108,215
Chyba stř. hodnoty	0,227	Chyba stř. hodnoty	0,302	Chyba stř. hodnoty	0,285
Medián	107,690	Medián	108,410	Medián	108,520
Modus	106,530	Modus	109,680	Modus	107,040
Směr. odchylka	1,873	Směr. odchylka	2,154	Směr. odchylka	1,890
Rozptyl výběru	3,508	Rozptyl výběru	4,638	Rozptyl výběru	3,572
Špičatost	-0,239	Špičatost	-0,131	Špičatost	-0,534
Šikmost	0,357	Šikmost	-0,268	Šikmost	-0,034
Rozdíl max-min	8,460	Rozdíl max-min	10,340	Rozdíl max-min	8,090
Minimum	104,740	Minimum	103,040	Minimum	104,480
Maximum	113,200	Maximum	113,380	Maximum	112,570
Součet	7338,99	Součet	5522,25	Součet	4761,48
Počet	0	Počet	0	Počet	0
Počet	68,000	Počet	51,000	Počet	44,000

Tabulka 12 Statistická data dle jednotlivých velikostí dávek 7/8 - pokračování

<i>191-200</i>		<i>201-210</i>		<i>211-220</i>	
Stř. hodnota	108,077	Stř. hodnota	107,087	Stř. hodnota	108,455
Chyba stř. hodnoty	0,258	Chyba stř. hodnoty	0,357	Chyba stř. hodnoty	0,286
Medián	107,900	Medián	107,250	Medián	108,520
Modus	109,210	Modus	není k dispozici	Modus	není k dispozici
Směr. odchylka	1,765	Směr. odchylka	1,892	Směr. odchylka	1,645
Rozptyl výběru	3,117	Rozptyl výběru	3,578	Rozptyl výběru	2,707
Špičatost	-0,702	Špičatost	-1,135	Špičatost	1,978
Šikmost	0,289	Šikmost	-0,016	Šikmost	-1,033
Rozdíl max-min	6,980	Rozdíl max-min	6,730	Rozdíl max-min	7,710
Minimum	105,050	Minimum	103,710	Minimum	103,330
Maximum	112,030	Maximum	110,440	Maximum	111,040
Součet	5079,600	Součet	2998,440	Součet	3579,010
Počet	47,000	Počet	28,000	Počet	33,000

Tabulka 13 Statistická data dle jednotlivých velikostí dávek 8/8- dokončení

<i>221-230</i>		<i>231-240</i>		<i>241-250</i>	
Stř. hodnota	107,620	Stř. hodnota	107,732	Stř. hodnota	107,994
Chyba stř. hodnoty	0,391	Chyba stř. hodnoty	0,397	Chyba stř. hodnoty	0,458
Medián	107,815	Medián	108,220	Medián	108,220
Modus	není k dispozici	Modus	není k dispozici	Modus	není k dispozici
Směr. odchylka	1,748	Směr. odchylka	2,064	Směr. odchylka	1,831
Rozptyl výběru	3,055	Rozptyl výběru	4,261	Rozptyl výběru	3,354
Špičatost	-0,753	Špičatost	-0,375	Špičatost	-0,871
Šikmost	0,112	Šikmost	0,059	Šikmost	-0,272
Rozdíl max-min	6,020	Rozdíl max-min	8,020	Rozdíl max-min	6,100
Minimum	104,710	Minimum	104,490	Minimum	104,600
Maximum	110,730	Maximum	112,510	Maximum	110,700
Součet	2152,400	Součet	2908,760	Součet	1727,910
Počet	20,000	Počet	27,000	Počet	16,000