

**Univerzita Pardubice**

**Fakulta ekonomicko-správní  
Ústav systémového inženýrství a informatiky**

**Využití autokorelační funkce při zpracování dat**

**Michaela Hettlerová**

**Diplomová práce  
2013**

Univerzita Pardubice  
Fakulta ekonomicko-správní  
Akademický rok: 2012/2013

## ZADÁNÍ DIPLOMOVÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: Bc. Michaela Knapová  
Osobní číslo: E10240  
Studijní program: N6209 Systémové inženýrství a informatika  
Studijní obor: Informatika ve veřejné správě  
Název tématu: Využití autokorelační funkce při zpracování dat  
Zadávající katedra: Ústav systémového inženýrství a informatiky

### Z á s a d y p r o v y p r a c o v á n í :

Cílem diplomové práce bude seznámení s danou problematikou. Bude vysvětlena korelační a autokorelační funkce. V práci bude ukázka využití autokorelační funkce v matematickém modelu. Model bude interpretován v prostředí MatLab a následně bude zhodnocen.

1. Uvedení do problematiky
2. Korelační a autokorelační funkce
3. Zpracování získaných dat
4. Verifikace získaného modelu s daty

Rozsah grafických prací:

Rozsah pracovní zprávy:

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

Seznam odborné literatury:

ANDĚL J.: Statistické metody. Praha : Matfyzpress, 2007. 299 s. ISBN 978-80-7378-003-6.

ČAPEK J., FABIÁN P.: Model of the object describing with help of the Weighting function. In. Processings of the 17th International Conference on CAD/CAM Robotics and Factoring of the Future, Durban, 2001 South Africa pp 1233-1238, ISBN 1-86840-432-3

ČAPEK J.: Multi-Input one Model Based on the Weighting Function. In. Proceedings of the 10th International DAAAM Symposium, DAAAM International, Editor B.Katalinic, Vienna, Austria, 1999. ISBN 3-901509-10-0.

HINDLS R., HRONOVÁ S., NOVÁK I.: Metody statistické analýzy pro ekonomy, Praha: Management Press, 2000. 259 s. ISBN 80-7261-013-9.

KARBAN P: Výpočty a simulace v programech Matlab a Simulink. Brno: Computer Press, 2006. 220 s. ISBN 80-251-1301-9.

KUBANOVÁ J.: Statistické metody pro ekonomickou a technickou praxi, Bratislava: Statis, 2003. 247 s. ISBN 80-85659-31-X.

TŮMA J.: Složité systémy řízení : 1. díl: Regulace soustav náhodnými poruchami. Ostrava: TECHNICKÁ UNIVERZITA OSTRAVA, 1997. 151 s. ISBN 80-7078-534-9.

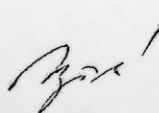
Vedoucí diplomové práce:

prof. Ing. Jan Čapek, CSc.


Ústav systémového inženýrství a informatiky

Datum zadání diplomové práce: **3. října 2012**

Termín odevzdání diplomové práce: **30. dubna 2013**

  
doc. Ing. Renáta Myšková, Ph.D.  
děkanka

L.S.

  
prof. Ing. Jan Čapek, CSc.  
vedoucí ústavu

V Pardubicích dne 3. října 2012

## **PROHLÁŠENÍ**

Prohlašuji, že jsem tuto práci vypracoval samostatně. Veškeré literární prameny a informace, které jsem v práci využil, jsou uvedeny v seznamu použité literatury.

Byl jsem seznámen s tím, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorský zákon, zejména se skutečností, že Univerzita Pardubice má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle § 60 odst. 1 autorského zákona, a s tím, že pokud dojde k užití této práce mnou nebo bude poskytnuta licence o užití jinému subjektu, je Univerzita Pardubice oprávněna ode mne požadovat přiměřený příspěvek na úhradu nákladů, které na vytvoření díla vynaložila, a to podle okolností až do jejich skutečné výše.

Souhlasím s prezenčním zpřístupněním své práce v Univerzitní knihovně.

V Pardubicích dne 15. 8. 2013

Bc. Michaela Hettlerová

## **PODĚKOVÁNÍ:**

Tímto bych ráda poděkovala svému vedoucímu práce prof. Ing. Janu Čapkovi, CSc. za jeho odbornou pomoc, cenné rady a poskytnuté materiály, které mi pomohly při zpracování diplomové práce.

## **ANOTACE**

*Tato práce se zabývá využitím autokorelační funkce při zpracování různých typů dat. Práce seznamuje s pojmy potřebnými k provádění analýz a modelování při využití autokorelační funkce. Popisuje práci s časovými řadami, signály a obrazy. Analýzy a modely popsané teoreticky jsou v práci ukázány i s konkrétními daty. Každý model je na závěr zhodnocen.*

## **KLÍČOVÁ SLOVA**

*Časové řady, analýza časových řad, korelační funkce, autokorelační funkce, parciální autokorelační funkce, signál, periodičita signálu, obraz, náhodný proces, Box-Jenkinsova metodologie*

## **TITLE**

*The use of autocorrelation functions for data processing*

## **ANNOTATION**

*This thesis deals with the use of the autocorrelation function when processing different types of data. The thesis introduces the concepts needed for the analysis and modeling using autocorrelation function. The thesis also describes work with time series, signals and images. Analysis and theoretical models described in this thesis use real data sources. Each model is evaluated at the end.*

## **KEYWORDS**

*Time series, analysis of time series, correlation function, autocorrelation function, partial autocorrelation function, the signal, frequency signal, image, random process, Box-Jenkins metodologie*

# OBSAH

ÚVOD .....	11
<b>1 UVEDENÍ DO PROBLEMATIKY .....</b>	<b>13</b>
1.1 ZÁKLADNÍ STATISTICKÉ CHARAKTERISTIKY .....	13
1.1.1 Střední hodnota .....	13
1.1.2 Rozptyl .....	14
1.1.3 Kovariance .....	14
1.1.4 Korelační koeficient .....	14
1.2 KORELAČNÍ A AUTOKORELAČNÍ FUNKCE .....	15
1.2.1 Korelační funkce .....	15
1.2.2 Křížově (vzájemně) korelační funkce .....	15
1.2.3 Autokorelační funkce .....	16
1.2.4 Parciální autokorelační funkce .....	17
1.3 BÍLÝ ŠUM .....	18
1.3.1 Proces bílého šumu .....	18
1.4 STOCHASTICKÝ PROCES .....	18
1.4.1 Stacionarita procesu .....	19
<b>2 ČASOVÉ ŘADY .....</b>	<b>21</b>
2.1 DRUHY ČASOVÝCH ŘAD .....	21
2.2 ZOBRAZENÍ ČASOVÝCH ŘAD .....	22
2.3 ANALÝZA ČASOVÝCH ŘAD .....	25
2.3.1 Dekompozice ČŘ .....	25
2.3.2 Boxova-Jenkinsova metologie .....	26
<b>3 SIGNÁL .....</b>	<b>33</b>
3.1 SIGNÁL A ČASOVÁ ŘADA .....	33
3.2 KLASIFIKACE SIGNÁLU .....	33
3.3 APLIKACE ACF A KORELAČNÍCH FUNKCÍ NA SIGNÁLY .....	34
3.3.1 Zjištění periodicity signálu .....	34
3.3.2 Detekce signálu v šumu .....	34
3.3.3 Nalezení časového zpoždění .....	34
3.3.4 Segmentace slov řečového signálu .....	35
<b>4 OBRAZ .....</b>	<b>36</b>
4.1 OBRAZ A NÁHODNÝ PROCES .....	36
4.2 KARHUNEN - LOEVE TRANSFORMACE OBRAZU .....	36
<b>5 NÁVRH MODELU .....</b>	<b>39</b>
5.1 MODELOVÁNÍ .....	39
5.1.1 Formulace problému .....	40
5.1.2 Základní návrh modelu .....	40
5.1.3 Implementace modelu .....	40
5.1.4 Verifikace a validace .....	40
5.1.5 Simulace a analýza .....	40
5.1.6 Sumarizace výsledku .....	40
5.2 INDEX SPOTŘEBITELSKÝCH CEN .....	41
5.3 MODEL PRO VOZOVÝ PARK .....	49
5.4 GENEROVÁNÍ SIGNÁLŮ .....	53
5.4.1 Zjištění periodicity signálu .....	53
5.4.2 Odstranění šumu z generovaného signálu .....	57
5.5 KARHUNEN-LOEVEHO TRANSFORMACE .....	61
5.5.1 Vliv komprese na kvalitu obrazu .....	61
5.5.1 Optimalizace procesu zpracování obrazu .....	66
5.5.2 Zpracování barevného obrazu .....	68
<b>ZÁVĚR .....</b>	<b>73</b>
<b>POUŽITÁ LITERATURA .....</b>	<b>75</b>

**SEZNAM TABULEK**

Tabulka 1: Souhrn průběhů teoretických ACF a PACF .....	31
Tabulka 2: Nalezení nejvýhodnějšího období .....	53

**SEZNAM GRAFŮ**

Graf 1: Průběh dat CPI .....	41
Graf 2: Zjištění mezí optLags .....	42
Graf 3: Průběh ACF a PACF .....	43
Graf 4: diferencovaná data .....	44
Graf 5: ACF a PACF po 1. diferenci .....	45
Graf 6: Srovnání ACF a PACF .....	46
Graf 7: Nekorelovanost reziduí .....	47
Graf 8: Q-Q graf reziduí .....	48
Graf 9: Průběh pohonných hmot .....	50
Graf 10: Zobrazení 1. difference .....	50
Graf 11: Hledání optimálního zpoždění .....	51
Graf 12: Průběh ACF pohonných hmot .....	52
Graf 13: Průběh generovaného signálu .....	55
Graf 14: Určení mezí optLags .....	56
Graf 15: Průběh ACF pro určení periodicity signálu .....	57
Graf 16: Vygenerovaný zašuměný signál .....	58
Graf 17: Zjištění parametru nLags .....	59
Graf 18: Vyčištěný signál .....	59
Graf 19: Srovnání zašuměného a vyčištěného signálu .....	60
Graf 20: Prostorové zobrazení ACF matice obrazu .....	62

**SEZNAM ILUSTRACÍ**

Obrázek 1: Stacionární (vlevo) a nestacionární řada (vpravo) .....	20
Obrázek 2: Ukázka spojnicového grafu .....	23
Obrázek 3: Ukázka sloupcového grafu .....	23
Obrázek 4: Graf ACF pro jednotlivá $k \geq 0$ .....	24
Obrázek 5: Určení časového zpoždění .....	35
Obrázek 6: Obraz jako náhodný proces .....	36
Obrázek 7: Zobrazení modelu .....	49
Obrázek 8: Srovnání komprese 40% a originálu .....	63
Obrázek 9: Srovnání komprese 44% a originálu .....	63
Obrázek 10: Srovnání komprese 48% a originálu .....	64
Obrázek 11: Srovnání komprese 52 % a originálu .....	64
Obrázek 12: Srovnání komprese 56 % a originálu .....	65
Obrázek 13: Srovnání komprese 80 % a originálu .....	65
Obrázek 14: Srovnání komprese 88 % a originálu .....	66
Obrázek 15: Vstupní obraz - vesmir.jpg .....	66
Obrázek 16: Původní obraz kalkulačky .....	68
Obrázek 17: Podobraz "red" – bez komprese .....	69
Obrázek 18: Podobraz "green" - bez komprese .....	70

Obrázek 19: Podobraz "blue" - bez komprese .....	70
Obrázek 20: Podobraz "red" - s kompresí .....	71
Obrázek 21: Podobraz "green" - s kompresí .....	71
Obrázek 22: Podobraz "blue" - s kompresí .....	72
Obrázek 23: Zkomprimovaný obraz .....	72

## SEZNAM ZKRATEK A ZNAČEK

ACF	Autokorelační funkce
ČŘ	Časová řada
KLT	Karhunen-Loeveho transformace
NS	Neuronová síť
PACF	Parciální autokorelační funkce

# ÚVOD

Jedním ze základních prostředků podávající informaci o charakteru stochastického procesu je autokorelační funkce. Empirická pozorování v různých oblastech jsou zpracovávány formou časových řad, buď jsou tak naměřeny nebo se převádějí do časových řad. Tato forma dat je lépe zpracovatelná a dá se s nimi provádět řada analýz, které usnadňují chápání jejich dynamiky vývoje.

V první kapitole jsou vysvětleny základní pojmy spojené s danou problematikou, které jsou v dalších kapitolách dále rozvinuty a použity. Vysvětluje statistické charakteristiky, které jsou dále použity k objasnění principu a formy korelační a autokorelační funkce.

Druhá kapitola pojednává o časových řadách, jak se s nimi dá pracovat. Dále obsahuje druhy časových řad, jejich dekompozici a analýzy zpracování dat. Hlavně se zabývá Box-Jenkinsovou metodologií, která při konstrukci modelu vychází z autokorelační a parciální autokorelační funkce.

Ve třetí kapitole je popsán další typ dat, který při zpracování využívá korelační analýzy a autokorelační funkci, tímto typem dat je signál, na který lze z určitého pohledu nahlížet, jako na časovou řadu. Kapitola popisuje, jak lze v této oblasti uplatnit autokorelační funkci.

Čtvrtá kapitola nejprve uvádí, jak lze nahlížet na obraz. Obraz popisuje, jako náhodný proces, při čemž může být v této formě zpracováván s využitím korelačních analýz a také autokorelační funkce. Hlavně se zaměřuje na Karhunen-Loeveho transformaci.

V páté kapitole je teoreticky nastíněno modelování, což je dále využito při konkrétním zpracování dat. Na časové řadě s daty udávajícími indexy spotřebitelských cen za jednotlivá čtvrtletí od roku 2000 až do roku 2013 je ukázáno hledání vhodného modelu řady dle Box-Jenkinsovy metodologie. Pro další data je ukázána případová studie, která modeluje situaci firmy s vozovým parkem hledající pomocí váhové funkce a autokorelační funkce odpověď na otázku, kdy se pro jejich firmu vyplatí nakoupit pohonné hmoty. Pro zpracování signálu je vygenerován signál, na němž jsou provedena pozorování dle teoretického popisu ve třetí kapitole. Pro detekci šumu v signálu je vygenerován další signál, který je zašuměn pomocí náhodné složky a tento šum je následně odstraněn za využití autokorelační funkce. Karhunen-Loeveho transformace je ukázána v rámci aplikace komprese obrazu, kterým je upravené černobílé logo Univerzity Pardubice, je na něm ukázán vliv komprese na kvalitu obrazu, jako další obraz je použit černobílý snímek vesmíru, kde je ukázána optimalizace

výpočtu a jako poslední obraz slouží barevná fotografie kalkulačky, kde je ukázáno navíc rozložení obrazu na jednotlivé barevné vrstvy RGB.

Vše je zpracováno za pomoci MATLAB R2010a, některé výstupy jsou upraveny v MS Excel.

**Cílem práce je analýza dat s využitím autokorelační funkce.**

# 1 UVEDENÍ DO PROBLEMATIKY

Tato kapitola se zabývá vysvětlením základních pojmů, které mají úzkou spojitost s danou problematikou. Pojmy nacházející se v této kapitole budou v dalších kapitolách rozšířeny a na jejich základě odvozeny jednotlivé funkce dané problematiky. Nejprve kapitola vysvětluje důležité statistické charakteristiky, z nichž vychází korelační a autokorelační funkce, což vysvětluje další podkapitola, která uvádí definice a zápisy korelační, křížově korelační, autokorelační a parciální autokorelační funkci. Dále kapitola popisuje bílý šum a jeho proces, jakožto určitý náhodný signál vykazující vlastnosti pomocí autokorelační funkce. Závěr kapitoly uvádí definici stochastického procesu a vysvětluje za jakých podmínek je tento proces stacionární.

## 1.1 Základní statistické charakteristiky

K tomu, aby mohla být v práci popsána autokorelační funkce, je potřeba znalosti základních statistických charakteristik, jako je střední hodnota, rozptyl, kovariance, korelační koeficient. Tato podkapitola vychází z publikací [1] a [2].

### 1.1.1 Střední hodnota

Střední hodnoty jsou určitá čísla zastupující hodnoty zkoumaného statistického znaku. Jedno číslo tak nahrazuje dlouhou řadu hodnot znaku, čímž usnadňuje porovnávání dvou i více statistických souborů.

Mezi střední hodnoty se řadí:

- **Aritmetický průměr** je podíl úhrnu hodnot proti rozsahu souboru, zapsaný ve vztahu (1)

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

Pokud se přidají váhy, tak se hovoří o **váženém** aritmetickém průměru.

- **Modus** je nejčastěji se vyskytující hodnota ve statistickém souboru.
- **Medián** je hodnota nacházející se ve středu souboru seřazeném podle velikosti

Pro práci s autokorelační funkcí se převážně vychází z následující definice střední hodnoty:

Nechť  $X$  je náhodná veličina s distribuční funkcí  $F(x)$ . Pak jako  $E(X)$  budeme značit střední hodnotu náhodné veličiny  $X$ . Vztahy (2) je střední hodnota pro diskrétní rozdělení s pravděpodobnostní funkcí  $P(x)$  a vztah (3) pro spojité rozdělení s hustotou  $f(x)$  [2].

$$E(x) = \sum_x xP(x) \quad (2)$$

$$E(x) = \int_{-\infty}^{\infty} xf(x)dx \quad (3)$$

Střední hodnota může být také označována jako  $\mu_x$ .

### 1.1.2 Rozptyl

Rozptyl udává míru variability náhodné veličiny. Dle zdroje [2] ho lze zapsat dle vztahu (4) pro diskrétní rozdělení s pravděpodobnostní funkcí  $P(x)$  a jako (5) pro spojité rozdělení s hustotou  $f(x)$ .

$$D(x) = \sum_x (x - E(X))^2 P(x) \quad (4)$$

$$D(x) = \int_{-\infty}^{\infty} (x - E(X))^2 f(x)dx \quad (5)$$

### 1.1.3 Kovariance

Kovariance dvou náhodných veličin  $X$  a  $Y$  se středními hodnotami  $\mu_x$  a  $\mu_y$  je dána vztahem (6) [2].

$$cov_{xy} = E(X - \mu_x)(Y - \mu_y) \quad (6)$$

### 1.1.4 Korelační koeficient

Korelační koeficient udává informaci o stupni závislosti dvou náhodných veličin. Platí, že  $D(X) > 0$  a  $D(Y) > 0$  a zápis je uveden ve vztahu (7) [3].

$$\rho_{xy} = \frac{cov_{xy}}{\sqrt{D(X)}^1 \sqrt{D(Y)}} \quad (7)$$

---

<sup>1</sup> Tento vztah lze označovat za směrodatnou odchylku.

## 1.2 Korelační a autokorelační funkce

Dle zdroje [3] je autokorelační funkce (ACF) a parciální autokorelační funkce (PACF) základním prostředkem podávající informaci o charakteru stochastického procesu. Při empirické práci s časovými řadami (ČŘ) je využíváno jejich výběrových protějšků. Každý z modelů třídy AR, MA a ARMA je charakteristický specifickou formou ACF a PACF, tudíž lze jejich odhady použít při identifikaci modelu konkrétních analyzovaných ČŘ. Při analýzách využívající ACF se ve většině případů uvažuje, že se jedná o stacionární data, popř. taková data, která lze na stacionární převést. Tato kapitola vysvětluje význam korelační a autokorelačních funkcí při zpracování dat.

### 1.2.1 Korelační funkce

Předpokladem jsou dvě náhodné veličiny  $x$  a  $y$ , které mají konečné nenulové rozptyly  $D(x)$  a  $D(y)$ . Pak se k vyjádření míry závislosti mezi těmito veličinami využívá následujícího vztahu (8) nazývaného korelační funkce [4].

$$\rho(x, y) = \frac{\text{cov}(x, y)}{\sqrt{D(x)}\sqrt{D(y)}} \quad (8)$$

kde  $D(x)$  je rozptyl veličiny  $x$ ,  $D(y)$  je rozptyl veličiny  $y$ . Zapisují se dle vztahů (9) a (10).

$$D(x) = E(x_t - \mu_t)^2 \quad (9)$$

$$D(y) = E(y_t - \mu_t)^2 \quad (10)$$

V případě, že jsou veličiny nekorelované (není mezi nimi korelační vztah), pak platí vztah (11).

$$\rho(x, y) = 0 \quad (11)$$

V opačném případě existuje mezi veličinami korelační vztah. Ačkoli jsou dvě náhodné veličiny nekorelované, nemusí být nezávislé.

### 1.2.2 Křížově (vzájemně) korelační funkce

Dle [5] lze pro dva slabě stacionární stochastické procesy definovat vztah (12) vzájemné korelační funkce.

$$\rho_k = \frac{\text{cov}(X_t, X_{t+k})}{\sqrt{D(X_t)}\sqrt{D(Y_t)}} = \frac{\gamma_k}{\sqrt{D(X_t)}\sqrt{D(Y_t)}} \quad (12)$$

Vzájemná korelační funkce je dle [6] a [7] reálná funkce časového posunutí. Pro dvě náhodné veličiny ji lze také zapsat, jako vztah (13)

$$\rho_{fg}(x_1, y_1; x_2, y_2) = E\{f(x_1, y_1) \cdot g(x_2, y_2)\} \quad (13)$$

### 1.2.3 Autokorelační funkce

Dle [8] autokorelační funkce (ACF) podává informaci o síle lineární závislosti mezi veličinami  $y_t, y_{t+k}$ . Korelace mezi náhodnými veličinami  $y_t$  a  $y_{t+k}$  však může být způsobena jejich korelací s veličinami  $y_{t-1}, y_{t-2}, \dots, x_{t+k+1}$ .

Lze tedy odvodit, že pokud se vztahuje korelační funkce sama k sobě, pak se jedná o autokorelační funkci ve tvaru (14).

$$\rho_k = \frac{cov(X_t, X_{t+k})}{D(X)} = \frac{\gamma_k}{\gamma_0} \quad (14)$$

$$\gamma_k = cov(x_t, x_{t+k}) = E[(x_t - \mu)(x_{t+k} - \mu)] \quad (15)$$

$$D(X_t) = D(X_{t+k}) = \gamma_0 \quad (16)$$

kde  $\gamma_k$  je autokovarianční funkce. Autokorelační funkce je symetrická kolem  $k=0$ ,  $\gamma_k = \gamma_{-k}$  a  $\rho_k = \rho_{-k}$ , proto je vyjadřována pouze pro  $k \geq 0$ . Nabývá hodnoty v intervalu  $\langle -1; 1 \rangle$ . Graf závislosti  $\rho_k$  na  $k$  se nazývá periodogram, přičemž platí  $\rho_0 = 1$  a  $|\rho_k| \leq 1$  pro  $k > 0$ . Graf autokorelační funkce se nazývá korelogram [9].

Dle literatury [6] hovoříme o autokorelační funkci, jestliže je hodnocena závislost hodnot jednoho náhodného procesu ve dvou různých časových okamžicích. V případě hodnocení závislosti dvou náhodných procesů ve dvou různých časových okamžicích, jedná se o křížovou, neboli vzájemnou korelační funkci (viz kapitola 1.2.2).

Souhrn vlastností ACF vyplývající v případě stacionárního stochastického procesu  $\{X_t\}$  je dle literatury [10] následující:

- ACF nultého řádu je rovna jedné ( $\rho_0 = 1$ )
- Absolutní hodnota autokovarianční funkce  $k$ -tého řádu je menší či rovna hodnotě autokovarianční funkce nultého řádu a absolutní hodnota ACF  $k$ -tého řádu je menší či rovna jedné ( $|\gamma_k| \leq \gamma_0; |\rho_k| \leq 1$ )
- ACF je pro všechna  $k$  symetrická kolem  $k=0$  ( $\gamma_k = \gamma_{-k}$  a  $\rho_k = \rho_{-k}$ ), což znamená, že jsou ACF sudé, z čehož vyplývá, že se lze omezit jen na  $k \geq 0$

Odvození lze nalézt v [11].

### 1.2.4 Parciální autokorelační funkce

Korelace mezi dvěmi náhodnými veličinami je velmi často způsobena působením korelace veličiny třetí nebo více veličinami. Parciální autokorelace nesou informaci o korelaci veličin  $X_t$  a  $X_{t-1}$  očištěné o vliv veličin mezi nimi. [10]

Parciální autokorelaci se zpožděním  $k$  určuje parciální regresní koeficient  $\phi_{kk}$  v autoregresi  $k$ -tého řádu

$$X_t = \phi_{k1}X_{t-1} + \phi_{k2}X_{t-2} + \dots + \phi_{kk}X_{t-k} + e_t$$

kde je  $e_t$  nekorelovaná s veličinami  $X_{t-j}$ ,  $j \geq 1$ . Po roznásobení a postupném použití Cramerova pravidla dostaneme parciální autokorelační funkci (PACF) ve tvaru (17) [12].

$$\phi_{kk} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_{k-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & 1 \end{vmatrix}} \quad (17)$$

Autokorelační funkce a parciální autokorelační funkce dané řady jsou v bodě  $k=1$  totožné a platí (18).

$$\rho_{11} = \rho_1 \quad (18)$$

### 1.3 Bílý šum

Ing. Tůma [6] definuje bílý šum, jako náhodný signál se zvláštními vlastnostmi. Ačkoliv je tento signál matematická abstrakce, lze pomocí něj modelovat náhodné chyby, jejichž základní charakteristikou je úplná nahodilost a žádná korelace mezi vzájemně časově posunutými hodnotami. Jiným uplatněním bílého šumu je při popisu přenosových vlastností lineárních dynamických systémů.

#### 1.3.1 Proces bílého šumu

Tento proces lze dle [13] charakterizovat jako stochastický proces  $\{\varepsilon_t\}$  řadou nekorelovaných náhodných veličin jednoho pravděpodobnostního rozdělení s konstantní střední hodnotou  $E(\varepsilon_t) = \mu_a$  (obvykle nulovou), konstantním rozptylem  $D(\varepsilon_t) = \sigma_a^2$  a  $\gamma_k = C(\varepsilon_t, \varepsilon_{t-k}) = 0$ , pro všechna  $k \neq 0$ . Tento proces je stacionární s autokorelační funkcí (ACF, viz kapitola 1.2.3):

$$\rho_k = \begin{cases} 1, & k = 0 \\ 0, & k \neq 0 \end{cases} \quad (19)$$

a parciální autokorelační funkcí (PACF, viz kapitola 1.2.4):

$$\phi_{kk} = \begin{cases} 1, & k = 0 \\ 0, & k \neq 0 \end{cases} \quad (20)$$

pak platí, že  $\rho_0 = \phi_{00} = 1$  pro jakýkoliv proces, pro  $k \neq 0$ .

### 1.4 Stochastický proces

Dle [3] je stochastický proces v čase uspořádaná řada náhodných veličin  $\{y(s,t), s \in S, t \in T\}$ , kde  $S$  je výběrový prostor a  $T$  je indexní řada. Pro každé  $t \in T$  je  $y(.,t)$  náhodná veličina definována na výběrovém prostoru  $S$ . Pro každé  $s \in S$  je  $y(s,.)$  realizace stochastického procesu definována na indexní řadě  $T$ , tj. uspořádaná řada čísel, z níž každé odpovídá jedné hodnotě indexní řady. Časovou řadu lze tedy chápat jako realizaci stochastického procesu.

Pro stochastický proces je vhodné dle [14] zavést následující vztahy odvozené ze vztahů (2),(3),(4),(5),(6) a (7):

- Střední hodnotu  $\mu_t = E(y_t)$

- Rozptyl  $\sigma^2 = D(y_t) = E(y_t - \mu_t)^2$
- Kovarianční funkci  $\gamma(t, t - k) = E(y_t - \mu_t)(y_{t-k} - \mu_{t-k})$
- Korelační funkci  $\rho(t, t + k) = \frac{\gamma(t, t+k)}{\sigma_t \sigma_{t+k}}$

### 1.4.1 Stacionarita procesu

Stochastický proces je označován jako stacionární, jestliže jsou charakteristiky jeho náhodných veličin v čase neměnné, tedy aby měl konstantní střední hodnotu, rozptyl a tudíž platilo (21) [12].

$$\rho_{(t,t+k)} = \gamma_{(t,t+k)} = \rho_k \quad (21)$$

Stacionarita je buď silná (striktní) nebo slabá (kovarianční). **Striktně stacionární** proces je popsán v [10], jako stochastický proces, u kterého platí, že jeho chování je časově invariantní, tj. pravděpodobnostní chování náhodného vektoru  $(y_{t1}, y_{t2}, \dots, y_{tk})$  je stejné jako rozdělení vektoru  $(y_{t1+h}, y_{t2+h}, \dots, y_{tk+h})$  pro libovolné  $h$ .

Platí-li, že  $\mu_t = \mu$  a  $\sigma_t^2 = \sigma^2$  pro všechna  $t$  a kovarianční a korelační funkce závisí jen na časové vzdálenosti náhodných veličin, pak se jedná o **slabě stacionární** stochastický proces. Pro tuto vlastnost platí vztah (22).

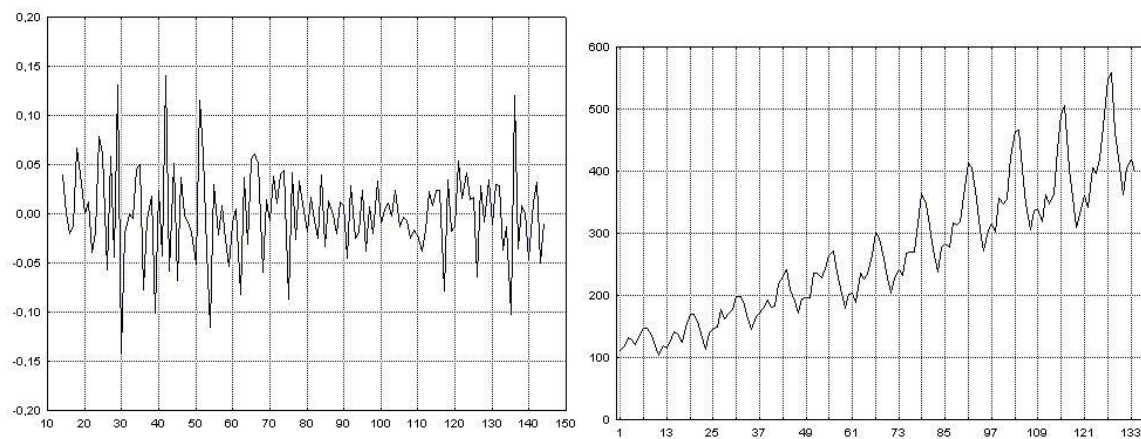
$$\text{cov}(y_t, y_s) = \text{cov}(y_{t+h}, y_{s+h}) \text{ pro libovolné } h \quad (22)$$

Závislost mezi dvěma libovolnými pozorování je jen v jejich vzájemné vzdálenosti a ne na skutečném časovém umístění v řadě.

Ve většině případů se předpokládá, že pokud proces není normální, tak má slabou stacionaritu, jelikož je méně omezující. Základní prostředek, který podává informaci o charakteru stochastického procesu je autokorelační a parciální autokorelační funkce. Pozorování těchto dvou funkcí umožňuje identifikovat konkrétní analyzované řady.

---

<sup>2</sup> Pokud neumocníme hovoříme o směrodatné odchylce ( $\sigma$ ).



**Obrázek 1:** Stacionární (vlevo) a nestacionární řada (vpravo)

*Zdroj [15]*

Obrázek 1 ukazuje typický průběh stacionární a nestacionární řady.

## 2 ČASOVÉ ŘADY

Obecně lze časové řady charakterizovat, jako každý jev, který je nějakým způsobem závislý na čase. Tyto jevy se lidé snaží popisovat a následně dle záznamů různými způsoby předpovídat jejich následný vývoj.

Časová řada (ČŘ) je v publikaci [75] charakterizována jako data, která jsou chronologicky uspořádána.

Dle [76] je časová řada posloupnost hodnot ukazatelů, měřených v určitých časových intervalech. Tyto intervaly jsou zpravidla rovnoměrné (ekvidistantní), a proto je můžeme zapsat následujícím způsobem:

$y_1, y_2, \dots, y_n$  neboli  $y_t, t=1, \dots, n$ , kde  $y$  značí analyzovaný ukazatel,  $t$  je časová proměnná s celkovým počtem pozorování  $n$ .

### 2.1 Druhy časových řad

ČŘ lze členit podle charakteru ukazatele na okamžikové a intervalové a podle druhu ukazatele na absolutní a odvozené [76].

**ČŘ intervalových ukazatelů** jsou naměřeny v závislosti na délce intervalu, za který je sledována, např. nezaměstnanost za daný rok. Je nutno dodržovat stejnou délku intervalů, při nedodržení intervalu je nutno provést přepočít (očistění) na jednotkový interval.

**Okamžikové ČŘ** se vztahují ke konkrétnímu časovému okamžiku, např. počet obyvatel k 31.12.

**Absolutní ČŘ** jsou dány již očištěnými ukazateli oproti tomu odvozené ČŘ jsou dány součty či poměrem.

V literatuře [76] jsou mimo uvedené ČŘ rozděleny podle:

- **periodicity sledování:**
  - **dlouhodobé ČŘ** – sledované období hodnot ukazatelů je rok nebo delší časový horizont, např. vývoj inflace za roky 2005-2012.
  - **krátkodobé ČŘ** – jsou za období kratší jak jeden rok, např. čtvrtletní hodnoty míry inflace za roky 2008-2010
  - **vysokofrekvenční ČŘ** – sestaveny z hodnot naměřené v časovém intervalu den, hodiny, minuty nebo sekundy, např. hodnoty měnového kurzu

- **druhu sledovaných ukazatelů:**
  - **primární ČŘ** – např. počet dokončených bytů za daný rok
  - **sekundární ČŘ** – např. roční zisk
- **homogenity ČŘ:**
  - **časové vymezení** – hodnoty se zjišťují za stejně dlouhé období
  - **prostorové vymezení** – měření se uskutečňuje na stanoveném území hranic
  - **věcné vymezení** – jasně definovaný ukazatel s danou měrnou jednotkou

## 2.2 Zobrazení časových řad

Nejčastějším zobrazením ČŘ jsou grafy. Tato forma zobrazení je přehledná a velmi snadno interpretovatelná. Toto zobrazení je důležité pro předběžnou analýzu dat, usnadňuje výběr modelu i přesnější výpočet a snadnější porozumění vztahům mezi daty. Data ČŘ jsou zobrazeny postupně v čase. Konstrukce grafů ČŘ může být považována za jednoduchou, ale je třeba brát v úvahu obecné principy označení os a proměnných, tak aby nedošlo ke zkreslení hodnot a následně ke klamavým informacím. Při aplikaci autokorelační funkce do analýz se převážně vychází právě z grafického vyjádření jejího průběhu. Členění grafů dle publikace [76]:

### **Grafy jedné a více ČŘ**

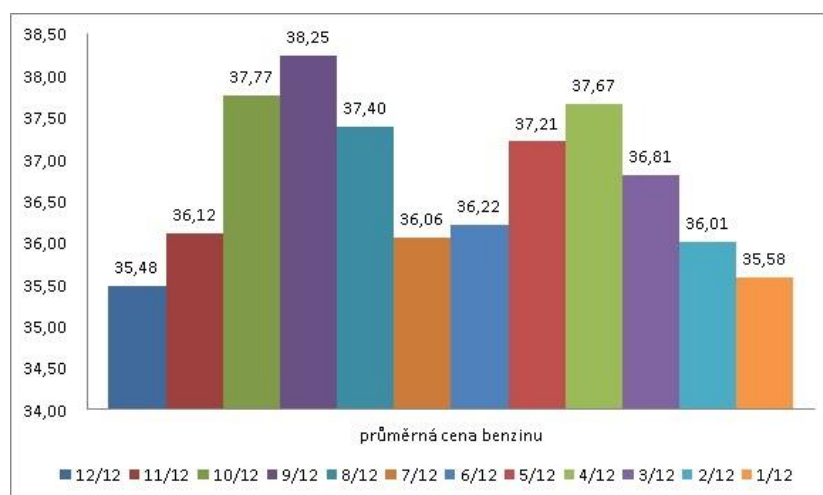
**Spojnicový graf** umožňuje získání prvotní informace o ČŘ, jednotlivé hodnoty ČŘ jsou zakresleny do souřadnicových os, horizontální osa zaznamenává časovou proměnnou a vertikální osa hodnoty ČŘ. Ukázka grafu na obrázku 2 vychází z dat získaných na [www.kurzy.cz](http://www.kurzy.cz).



**Obrázek 2:** Ukázka spojnicového grafu

*Zdroj [vlastní]*

**Sloupkový graf** je nejvíce využíván pro vyjádření vývoje ekonomických ukazatelů, tento typ grafu má několik modifikací. V ukázce na obrázku 3 jsou zobrazeny průměrné měsíční ceny benzínu za rok 2012 také na datech získaných na [www.kurzy.cz](http://www.kurzy.cz).



**Obrázek 3:** Ukázka sloupcového grafu

*Zdroj [vlastní]*

**Plošný graf** je modifikací grafu spojnicového, lze využít pro vyjádření vývoje jedné nebo více ČŘ. Ukázka opět vychází z dat, jako u spojnicového grafu.

**Graf „seasonal subseries“** je využíván při analýze sezónních ČŘ, data jsou v něm uspořádány podle jednotlivých sezón, např. zobrazení odlišnosti hodnot daného měsíce od průměrné hodnoty tohoto měsíce.

**Bodový diagram** – je nejjednodušší, nejsrozumitelnější a nejvíce používaný pro zobrazení vztahu dvou ČŘ (dvourozměrný bodový diagram).

**Matice bodových diagramů** – v případě, že je potřeba porovnat více dvojic ČŘ, tak se používá matice bodových diagramů.

### Grafy průzkumové analýzy ČŘ

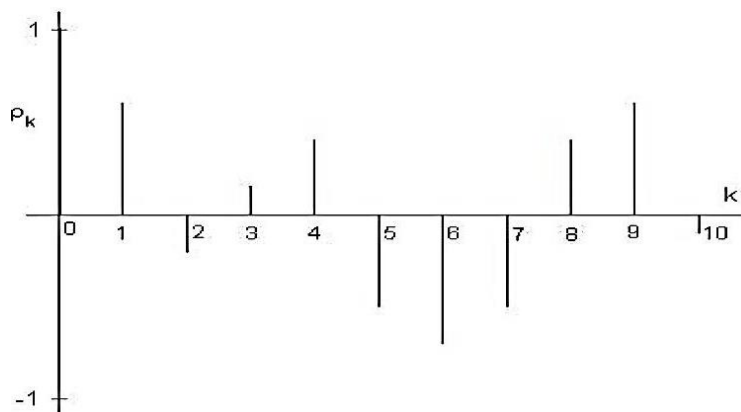
**Krabičkový graf („boxplot“)** má základním prvkem krabičku, která tvoří mezikvartilové rozpětí  $\tilde{x}_{75} - \tilde{x}_{25}$ , tudíž leží v krabičce 50% hodnot ČŘ, uvnitř je vodorovnou čarou vyznačen medián  $\tilde{x}_{50}$  a křížek značí průměr.

**Číslíkový histogram** udává v prvním sloupci hodnoty kumulovaných četností příslušející jednotlivým desítkám (stovkám) až do skupiny obsahující medián.

### Grafy pro identifikaci, testování a modelování složek ČŘ

**Histogram** vyjadřuje intervalové rozdělení četností řady, používá se k předběžnému posouzení vlastností rozdělení četností a porovnání s jiným rozdělením četností.

**Korelogram** ACF a PACF se používá především pro posouzení, zda má řada reziduí charakter bílého šumu. Ukázka korelogramu je na obrázku 4.



**Obrázek 4:** Graf ACF pro jednotlivá  $k \geq 0$

*Zdroj [19]*

**Periodogram** se využívá především, jako prostředek analýzy sezónních ČŘ a pro vyhledávání významných periodických složek v ČŘ.

**Kumulovaný periodogram** je také pro analýzu sezónních ČŘ, často se využívá ve spojení se sezónním očišťováním, s jeho pomocí se posuzuje, jestli odhad reziduální složky obsahuje nějakou periodickou složku nebo má charakter bílého šumu, čímž se zkoumá kvalita provedeného sezónního očištění.

## 2.3 Analýza časových řad

Analýzy ČŘ mají za cíl ve většině případů konstrukci vhodného modelu, což umožňuje porozumět mechanismu, díky němuž vznikají hodnoty časové řady a pochopení podmínek a vazeb, které způsobují vznik těchto hodnot. Změny podmínek a vazeb umožňují simulaci vlivu působícího změny ve vývoji ČŘ. Analýza bývá prováděna také z důvodu předpovědi budoucího chování. Mezi nejčastěji využívané metody analýz ČŘ se dle [76] zařazují následující:

- **Dekompozice ČŘ:** principem je rozložení ČŘ na čtyři základní složky, blíže v kapitole 2.3.1
- **Boxova-Jenkinsonova metodologie:** lépe se adaptují na změny v průběhu ČŘ a základním matematickým nástrojem pro analýzu ČŘ jsou v tomto případě metody korelační analýzy, tudíž využívá při práci autokorelační funkci a jiné korelační funkce, problematiku popisuje kapitola 2.3.2
- **Lineární kauzální (faktorové) modely:** hodnoty sledované ČŘ jsou v tomto modelu vysvětlovány pomocí jiných, tzv. faktorových ČŘ
- **Spektrální analýza ČŘ:** ČŘ je v tomto přístupu brána, jako nekoneční lineární kombinace sinusových a kosinusových funkcí s různými amplitudami a frekvencemi, intenzitu zastoupení jednotlivých frekvencí lze získat např. z periodogramu (viz kapitola 2.2), především je využívána ve Fourierově analýze

### 2.3.1 Dekompozice ČŘ

Každá ČŘ může obsahovat čtyři složky a dá se zapsat formou aditivního modelu ve tvaru (23) [20].

$$y_t = Tr_t + C_t + S_t + \varepsilon_t \quad (23)$$

Kde  $Tr_t$  je trendová složka,  $C_t$  je složka cyklická,  $S_t$  složka sezónní a  $\varepsilon_t$  je náhodná složka. Rozklad neboli dekompozice ČŘ na zmíněné složky má za cíl usnadnit identifikaci pravidelného chování ČŘ než původní nerozložené ČŘ [16].

ČŘ chápeme jako jistý trend, na který se „nabalují“ sezónní a cyklické složky a náhodná složka, kterou bývá převážně bílý šum [8].

**Trend** je výsledkem dlouhodobých a stálých procesů a ukazuje obecnou tendenci vývoje zkoumaného jevu za dlouhé období. Trend bývá převážně modelován pomocí matematických křivek. Existují i ČŘ bez trendu.

**Sezónní složka** je jistá odchylka od trendu, která se opakuje. Perioda této složky je menší než celkové sledované období

**Cyklická složka** udává kolísání okolo trendu, což je způsobeno v důsledku dlouhodobého cyklického vývoje, střídají se tu fáze růstu a poklesu.

**Náhodná složka** (též označovaná stochastická) zobrazuje nahodilé a jiné nesystematické výkyvy. Předpokladem je, že je tato složka tvořena tzv. bílým šumem (viz kapitola 1.3) s normálním rozdělením.

### 2.3.2 Boxova-Jenkinsova metodologie

Box- Jenkinsova metodologie se zabývá stavbou modelu ČŘ, kde je základním prvkem reziduální složka vykazující korelované náhodné veličiny. Tato metodologie tedy může zpracovávat ČŘ s navzájem závislými pozorováními. Vyžaduje určitou minimální délku alespoň 50 pozorování. Hlavním prostředkem výpočtů je využití korelační analýzy. Analýza se tu provádí systematicky dle následujících bodů:

- identifikace modelu
- odhad parametrů
- ověřování modelu

Dle [8] se u této metodologie u většiny případů předpokládá slabá stacionarita ČŘ nebo, že lze ČŘ převést na stacionární. Základním prvkem při konstrukci modelu je reziduální (náhodná) složka. Mezi typické modely patří specifické případy lineárních modelů MA, AR a ARMA.

V literatuře [5] je popisována jako metodologie analýzy ČŘ, které chápeme, jako realizaci náhodného (stochastického) procesu, kde klíčovou roli hrají autoregresní procesy a procesy klouzavých součtů.

#### Lineární proces

Je nekonečná řada, kterou lze dle [8] zapsat ve tvaru (24).

$$y_t = \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots \quad (24)$$

kde  $\psi_i$  jsou nějaké parametry a  $\varepsilon_t$  je vzájemně nekorelovaná a identicky rozdělená složka reprezentující tzv. bílý šum s nulovou střední hodnotou a rozptylem. Existence lineárního procesu je podmíněna konvergencí nekonečné řady náhodných veličin na pravé straně uvedeného vztahu.

### Autoregresní proces p-tého řádu AR(p)

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t \quad (25)$$

kde  $\phi_1, \phi_2, \dots, \phi_p$  jsou nenulové parametry. Při zavedení operátoru zpětného posunutí  $B^j X_t = X_{t-j}$  dostaneme AR(p) ve tvaru:

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) X_t = \phi_p(B) X_t = \varepsilon_t \quad (26)$$

kde  $\phi_p(B)$  je autoregresní operátor.

Autoregresní proces prvního řádu AR(1) lze odvodit do následujícího tvaru:

$$X_t = \phi X_{t-1} + \varepsilon_t \quad (27)$$

$$(1 - \phi B) X_t = \varepsilon_t \quad (28)$$

Jestliže pro AR(1) platí, že  $|\phi| < 1$ , pak je tento proces stacionární.

### **Vlastnosti tohoto procesu:**

- Nulová střední hodnota  $E(X_t) = 0$
- Autokorelační funkce  $\rho_k = \phi^k$
- Parciální autokorelační funkce  $\rho_{kk} = \phi$  pro  $k=1$   
 $\rho_{kk} = 0$  pro  $k>1$
- Rozptyl  $D(X_t) = \frac{\sigma^2}{1-\phi^2}$

### Procesy klouzavých součtů q-tého řádu MA(q)

Tvar modelu je následující:

$$X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (29)$$

kde  $\theta_1, \theta_2, \dots, \theta_q$  jsou nenulové parametry. Po zavedení operátoru klouzavých součtů  $\theta_q(B)$  dostaneme:

$$X_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \varepsilon_t = \theta_q(B) \varepsilon_t \quad (30)$$

Tento proces je vždy stacionární.

**Proces klouzavých součtů prvního řádu** MA(1) lze zapsat takto:

$$X_t = \varepsilon_t - \theta \varepsilon_{t-1} \quad (31)$$

**Vlastnosti tohoto procesu:**

- Nulová střední hodnota  $E(X_t) = 0$
- Autokorelační funkce  $\rho_k = -\frac{\theta}{1+\theta^2}$  pro  $k=1$
- $\rho_k = 0$  pro  $k>1$
- Parciální autokorelační funkce  $\rho_{kk} = \frac{\theta^k(1-\theta^k)}{1-\theta^{2(k+1)}}$
- Rozptyl  $D(X_t) = (1 + \theta^2)\sigma^2$

ACF je v tomto případě vhodným identifikačním nástrojem.

### **Smišené procesy ARMA**

ARMA je smíšený proces kombinací procesů AR a MA.

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (32)$$

Po zavedení operátoru má tento tvar:

$$\phi_p(B)X_t = \theta_q(B)\varepsilon_t \quad (33)$$

**Smišený proces prvního řádu** ARMA(1,1) má pak tento tvar:

$$X_t = \phi_1 X_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1} \quad (34)$$

Pokud platí, že  $|\phi| < 1$ , pak se jedná o proces stacionární (homogenní).

**Vlastnosti smíšeného procesu:**

- Nulová střední hodnota  $E(X_t) = 0$
- Rozptyl pro ARMA(1,1) je jedna  $D(X_t) = 1$
- Autokorelační funkce  $\rho_k = \frac{(1-\phi\theta)(\phi-\theta)}{1+\phi^2-2\phi\theta}$  pro  $k=1$
- $\rho_k = \phi \rho_{k-1}$  pro  $k>1$

Podobný průběh má i PACF, ale vyjádření je komplikované, tak nebývá uváděno.

## Kroky modelu

Zde jsou popsány jednotlivé kroky konstrukce Box-Jenkinsovy metodologie. Kroky dle autorů publikací [8] a [9] jsou následující:

1. **Identifikace modelu** má za úkol výběr vhodného modelu (MA, AR, ARMA) a určení řádu vybraného modelu. Postup může být následující:
  - a. Nejprve se alespoň vizuálně ověří, zda se jedná o již zmiňovanou stacionární řadu
  - b. Pokud se nejedná o stacionární řadu je potřeba ji stacionarizovat
  - c. V případě, kdy je střední hodnota stacionárního procesu nenulová, je potřeba řadu vycentrovat tzn., že bude od původní hodnoty odečten aritmetický průměr
  - d. Určení identifikačního bodu  $k_0$  autokorelační a parciální autokorelační funkce, který je v okamžiku, kdy začínají být obě funkce nulové, což se provádí pomocí odhadů (teoretických hodnot) ACF  $r_k$  a PACF  $r_{kk}$ , k čemuž je třeba určit také odhad autokovarianční funkce ve tvaru (36), jako střední hodnotu se uvažuje vztah ve tvaru (35), kde  $y_1, y_2, \dots, y_n$  je pozorovaná řada. Odhad ACF je dán vztahem (37) a odhad PACF je získán dosazením  $r_k$  do vztahu (17) za jednotlivé  $\rho_k$ , což lze uvést vztahem (38).

$$\bar{y} = \frac{1}{n} \sum_{t=1}^n y_t \quad (35)$$

$$c_k = \sum_{t=1}^{n-k} \frac{(y_t - \bar{y})(y_{t+k} - \bar{y})}{n} \quad \dots k=0, 1, \dots, n-1 \quad (36)$$

$$r_k = \frac{c_k}{c_0} \quad (37)$$

$$r_{kk} = \frac{r_k - \sum_{j=1}^{k-1} r_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j} r_{k-j}} \quad (38)$$

$$\text{kde } r_{kk} = r_{k-1,j} - r_{kk} r_{k-1,k-j} \text{ pro } j=1, 2, \dots, k-1$$

### **Zjištění nulovosti ACF a PACF**

O nulovosti autokorelační funkce se rozhoduje na základě Batlettovy aproximace [21], v případě parciální autokorelační funkce se rozhoduje pomocí Quenouilleovy aproximace

[22]. Je třeba zjistit, jak blízko nule musí být odhady ACF a PACF, aby se mohlo tvrdit, že  $\rho_k = 0$ , popř.  $\rho_{kk} = 0$ .

Bartlettova aproximace pro zjištění nulovosti ACF říká, že je-li  $\rho_k = 0$  pro  $k > k_0$ , pak platí (39).

$$\sigma(r_k) = \sqrt{\text{var}(r_k)} \approx \sqrt{\frac{1}{n} \left( 1 + 2 \sum_{j=1}^{k_0} r_j^2 \right)}, \quad k > k_0 \quad (39)$$

Testuje se  $H_0: \rho_k = 0$  proti  $H_1: \rho_k \neq 0$ . Je-li  $|r_k| > 2 \cdot \sigma(r_k)$ , pak  $H_0$  zamítáme na hladině spolehlivosti 0,05.

O nulovosti PACF se rozhoduje Quenouilleovou aproximací, je-li  $\rho_{kk} = 0$  pro  $k > k_0$ , pak platí (40).

$$\sigma(r_{kk}) = \sqrt{\text{var}(r_{kk})} \approx \sqrt{\frac{1}{n}}, \quad k > k_0 \quad (40)$$

Testuje se  $H_0: \rho_{kk} = 0$  proti  $H_1: \rho_{kk} \neq 0$ . Je-li  $|r_{kk}| > 2 \cdot \sigma(r_{kk})$ , pak  $H_0$  zamítáme na hladině spolehlivosti 0,05.

Průběh teoretických hodnot ACF a PACF sloužící k lepší identifikaci modelu je uveden v tabulce 1.

**Tabulka 1:** Souhrn průběhů teoretických ACF a PACF

<b>Model</b>	<b>ACF</b>	<b>PACF</b>
AR (1)	Exponenciálně nebo oscilačně klesá	Nulová pro $k > 1$
AR (2)	Exponenciálně nebo oscilačně klesá	Nulová pro $k > 2$
AR (p)	Exponenciálně nebo oscilačně klesá	Nulová pro $k > p$
MA (1)	Nulová pro $k > 1$	Exponenciálně nebo oscilačně klesá
MA (2)	Nulová pro $k > 2$	Exponenciálně nebo oscilačně klesá
MA (q)	Nulová pro $k > q$	Exponenciálně nebo oscilačně klesá
ARMA (1,1)	Exponenciálně nebo oscilačně klesá pro $k \geq 1$	Exponenciálně nebo oscilačně klesá
ARMA (p,q)	Po q posunutích jako pro A (p)	Po p posunutích jako pro MA (q)

*Zdroj: [5]*

**2. Odhad parametrů modelu** se provádí za pomoci následujících metod pro odhady parametrů:

- a. Metoda nejmenších nelineárních čtverců
- b. Podmíněná metoda nejmenších nelineárních čtverců
- c. Nepodmíněná metoda nejmenších nelineárních čtverců

Cílem této etapy je určení přesnosti získaných odhadů parametrů modelů AR, MA a ARMA.

**3. Verifikace modelu** má za úkol potvrzení správnosti a pravdivosti navrženého modelu. Pro verifikaci modelu bývají nejčastěji dle [8] používány následující metody:

- a. Metoda přeparametrizování modelu

- b. Metoda odhadnutých reziduí
- c. Metoda založená na portmanteau testu

V případě nestacionárních dat je potřeba data stacionarizovat, což bývá zpravidla u finančních časových dat, a to díky časově proměnné střední hodnotě nebo časově proměnného rozptylu. Tyto modely, tak lze ve většině případů modelovat pomocí modelů ARIMA, které vedle náhodných fluktuací modelují i trendovou složku. Stacionarita se v mnoha případech docílí diferencováním původní ČŘ. Jakmile se z nestacionární řady stane řada stacionární lze již proces vyjádřit pomocí modelů ARMA.

Podle [15] má tato metodologie jisté výhody, ale i nevýhody. Výhodou může být flexibilita a rychlost adaptace na změnu v charakteru modelovaného procesu a v mnoha případech dává lepší výsledky. Nevýhodou je, musí být dostatečně dlouhá realizace a ztrácí se možnost jednoduché interpretace výsledných modelů.

## 3 SIGNÁL

Signál lze obecně popsat, jako průběh determinované nebo náhodné fyzikální veličiny. Se signálem se v dnešní době setká každý, protože každá domácnost využívá zařízení, jež přijímají nějaký signál a díky němu jsou přenášeny informace, data, obraz, zvuk atd. V této kapitole bude kladen důraz na seznámení se s hlavními úlohami ACF při zpracování signálů.

### 3.1 Signál a časová řada

V publikacích [23] a [11] je ČŘ přiřazována svým charakterem ke skupině diskretních signálů. Na základě tohoto zjištění lze tvrdit, že se při zpracování signálů mohou používat stejné metody, jako u zpracování ČŘ (dekompozice, Box-Jenkinson apod.).

### 3.2 Klasifikace signálu

Signál je v literatuře [24] obecně charakterizován, jako fyzikální jev, chemické, biologické, ekonomické nebo jiné materiální povahy, jež nese informaci o stavu systému, který jej generuje. Signály lze klasifikovat, jako:

- spojité a diskretní
- reálné a komplexní
- deterministické a náhodné
- sudé a liché
- periodické a neperiodické
- jednorozměrné a vícerozměrné

### 3.3 Aplikace ACF a korelačních funkcí na signály

Při zpracování signálů se využívají korelační a autokorelační funkce zejména pro [25]:

- zjištění periodicity signálu, protože ACF je periodická se stejnou periodou, jako je signál
- detekci signálu v šumu
- nalezení časového zpoždění
- segmentaci slov řečového signálu

#### 3.3.1 Zjištění periodicity signálu

Pro zjištění periodicity signálu je důležité vycházet z grafického průběhu ACF. Graf zaznamenává průběh ACF ve tvaru (41), kde  $T$  je shodná periodičita pro dva periodické průběhy. Po zobrazení průběhu ACF je snadné periodicitu vyčíst, jelikož má ACF stejnou periodu, jako původní signál.

$$\rho_x(\tau) = \frac{1}{T} \int x(t)x(t + \tau)dt \quad (41)$$

#### 3.3.2 Detekce signálu v šumu

Autoři Pavlík a Poláček uvádějí ve své publikaci [26], že lze pro detekci signálu v šumu využít přímo autokorelační funkci tohoto signálu, která je dána střední hodnotou součinu originální a posunuté posloupnosti vstupního signálu, tvar zachycuje (42), kde se jako  $x(n)$  uvažuje originální posloupnost signálu a  $x(n+m)$  je posunutá posloupnost vstupního signálu.

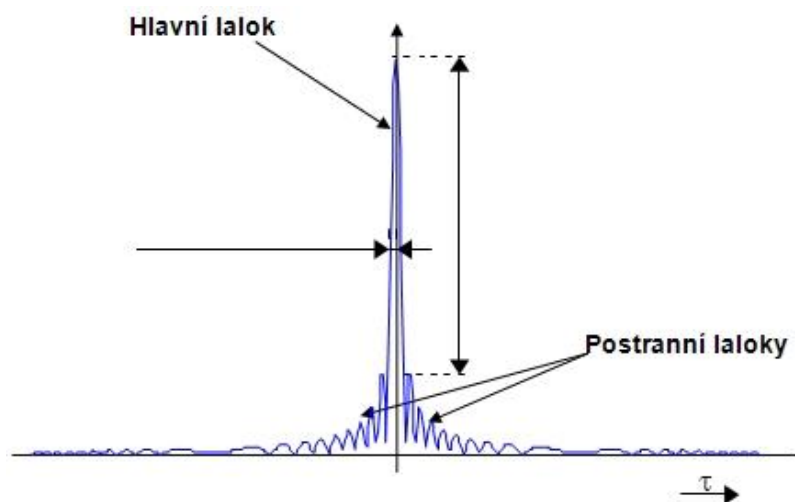
$$\rho_{rr}(m) = \frac{1}{N} \sum_{n=1}^{N-|m|} (r(n)) (r(n + m)), m \in (-N, N), n = 1, \dots, N \quad (42)$$

kde  $N$  označuje počet vzorků segmentu analyzovaného signálu.

#### 3.3.3 Nalezení časového zpoždění

K nalezení časového zpoždění signálu  $x(t)$  se využívá absolutní maximum autokorelační funkce, které se nachází v nejvyšším bodě grafu průběhu ACF. Autokorelační funkci tohoto signálu lze zapsat ve tvaru (43) [27].

$$\rho_x(\tau) = \int_{-\infty}^{\infty} x(t)x(t + \tau)dt \quad (43)$$



**Obrázek 5:** Určení časového zpoždění

*Zdroj [28]*

Obrázek 5 znázorňuje proces určení časového zpoždění, kdy se za časové zpoždění považuje špička hlavního laloku.

### 3.3.4 Segmentace slov řečového signálu

Autokorelační funkce najde většinu hranic, ale nemusí tomu tak vždycky být. Při segmentaci slov řečového signálu se vychází z krátkodobé autokorelační funkce, kterou lze dle [29] zapsat ve tvaru (44).

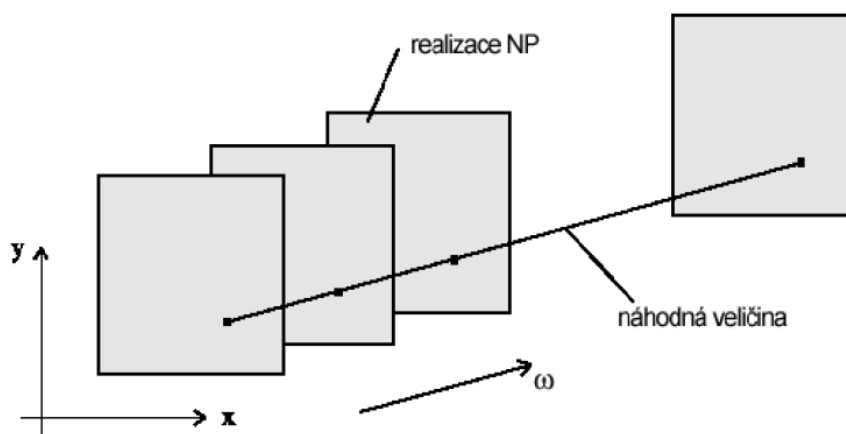
$$R_{xx}(m) = \sum_{n=0}^{N-1} x(n)x(n; m) \quad (44)$$

## 4 OBRAZ

Obraz lze chápat jako jakékoliv grafické vyjádření, ať už je to zachycení určitého předmětu, tváře, významné události apod., to vše lze poté nazvat obrazem. V této kapitole budeme obraz chápat, jako náhodný proces.

### 4.1 Obraz a náhodný proces

Při zpracování obrazu s využitím autokorelační funkce je důležité uvažovat obraz, jako náhodný proces. Dle zdroje [30] a [31] uvažujeme předpoklad vícenásobného snímání obrazu, kdy tak získáme soubor obrázků, jež nazýváme náhodným procesem. Jednotlivé realizace náhodného procesu považujeme za funkci  $f(x,y,\omega_i)$ , kde  $x$  a  $y$  určují polohu náhodných veličin náhodného procesu v prostorových souřadnicích a  $\omega_i, i=1, \dots, k$  udává polohu  $i$ -tého obrazu v náhodném procesu. Pak pro pevné  $[x,y]$  a proměnné  $i$  dostaneme jednu náhodnou veličinu a pro pevné  $i$  a proměnné hodnoty  $[x,y]$  získáme jeden obraz (viz obrázek 6).



Obrázek 6: Obraz jako náhodný proces

Zdroj: [30]

### 4.2 Karhunen - Loeve transformace obrazu

Tato kapitola vychází z publikací [30], [32] a [33]. Karhunen-Loeveho transformace (KLT) je matematický nástroj s velkým významem v širokém spektru aplikací. Především je využívána pro zpracování obrazové informace. Mezi hlavní úlohy KLT patří komprese obrazu. Nevýhodou této transformace je její vysoká výpočetní náročnost, tudíž i paměťová a časová náročnost. Transformace je založena na numericky vypočtených vlastních číslech a vlastních vektorů příslušné kovarianční matice. Cílem KLT je snížení dimenze prostoru

vstupních dat. Hledají se hlavní komponenty, které tvoří výběrový rozptyl a vektor koeficientů. Hlavním požadavkem je vzájemná nekorelovanost koeficientů hlavních komponent. Výběrový rozptyl je roven vlastnímu číslu a vektor koeficientů je rovný vlastnímu vektoru kovarianční matice zkonstruované ze vstupních dat.

Obraz, jak již bylo uvedeno, je považován za náhodný proces. Statistické charakteristiky tohoto náhodného procesu jsou následující:

**Střední hodnota** pro jednu náhodnou veličinu na pozici  $[x,y]$  v náhodném procesu, vztah (45)

$$\mu_f(x,y) = E\{f(x,y)\} = \int_{-\infty}^{\infty} z \cdot p(z;x,y) dz \quad (45)$$

kde  $z$  označuje hodnoty náhodné veličiny  $f(x,y,\omega_i)$ ,  $i=1,2,\dots,k$ .

**Autokorelační funkce** dvou různých náhodných veličin na pozicích  $[x_1, y_1]$ ,  $[x_2, y_2]$  z jednoho náhodného procesu je zapsaná ve tvaru (46).

$$\rho_{ff}(x_1, y_1; x_2, y_2) = E\{f(x_1, y_1, \omega_i) \cdot f(x_2, y_2, \omega_i)\} \quad (46)$$

Z jednotlivých hodnot ACF lze sestavit **autokorelační matice** (47). Na hlavní diagonále se nachází rozptyly jednotlivých náhodných veličin náhodného procesu.

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \rightarrow \begin{pmatrix} aa & ad & ag & ab & ae & ah & ac & af & ai \\ da & dd & dg & db & de & dh & dc & df & di \\ ga & gd & gg & gb & ge & gh & gc & gf & gi \\ ba & bd & bg & bb & be & bh & bc & bf & bi \\ ea & ed & eg & eb & ee & eh & ec & ef & ei \\ ha & hd & hg & hb & he & hh & hc & hf & hi \\ ca & cd & cg & cb & ce & ch & cc & cf & ci \\ fa & fd & fg & fb & fe & fh & fc & ff & fi \\ ia & id & ig & ib & ie & ih & ic & if & ii \end{pmatrix} \quad (47)$$

Korelační dvojice v uvedené matici byly vytvořeny dle [skripta.pdf] postupně po sloupcích, písmena  $a$  až  $i$  označují jednotlivé náhodné veličiny náhodného procesu.

**Autokovarianční funkce** dvou různých náhodných veličin na pozicích  $[x_1, y_1]$ ,  $[x_2, y_2]$  z jednoho náhodného procesu lze zapisovat ve tvaru (48).

$$\gamma_{ff}(x_1, y_1; x_2, y_2) = E\{[f(x_1, y_1, \omega_i) - \mu_f(x_1, y_1)] \cdot [f(x_2, y_2, \omega_i) - \mu_f(x_2, y_2)]\} \quad (48)$$

**Vzájemná korelace** dvou různých náhodných veličin na pozicích  $[x_1, y_1]$ ,  $[x_2, y_2]$  z jednoho náhodného je ve tvaru (49).

$$\rho_{fg}(x_1, y_1, \omega_i; x_2, y_2, \omega_j) = E\{f(x_1, y_1, \omega_i) \cdot g(x_2, y_2, \omega_j)\} \quad (49)$$

**Vzájemnou kovarianci** lze vyjádřit obdobně, a to následujícím vztahem (50).

$$\gamma_{fg}(x_1, y_1, \omega_i; x_2, y_2, \omega_j) = E\{[f(x_1, y_1, \omega_i) - \mu_f(x_1, y_1, \omega_i)] \cdot [g(x_2, y_2, \omega_j) - \mu_g(x_2, y_2, \omega_j)]\} \quad (50)$$

Pokud platí, že  $\gamma_{fg} = 0$ , pak jsou dva náhodné procesy nekorelované a platí (51).

$$E\{f(x_1, y_1, \omega_i) \cdot g(x_2, y_2, \omega_j)\} = E\{f(x_1, y_1, \omega_i)\} \cdot E\{g(x_2, y_2, \omega_j)\} \quad (51)$$

### **Kroky KLT mohou být následující:**

1. Z obrazu  $g$  se vytvoří průměr všech pixelů  $\mu_g$
2. Tento průměr se odečte od matice obrazu
3. Vytvoří se autokorelační matice nové matice z 2. kroku  $g - \mu_g$
4. Vypočtou se vlastní čísla a vlastní vektory této autokorelační matice
5. Vlastní čísla tvoří hlavní diagonálu matice vlastních čísel značenou  $C_{\tilde{g}\tilde{g}}$
6. Vlastní vektory srovnané po řádcích tvoří transformační matici  $A$ , která v rozepsaném tvaru vypadá následovně (52)

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1N} & a_{1N+1} & \dots & a_{12N} & \dots & \dots & a_{1N^2} \\ a_{21} & a_{22} & \dots & a_{2N} & a_{2N+1} & \dots & a_{22N} & \dots & \dots & a_{2N^2} \\ \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots & \dots & \dots & \vdots \\ \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots & \dots & \dots & \vdots \\ a_{N1} & a_{N2} & \dots & a_{NN} & a_{NN+1} & \dots & a_{N2N} & \dots & \dots & a_{NN^2} \\ a_{N+11} & a_{N+12} & \dots & a_{N+1N} & a_{N+1N+1} & \dots & a_{N+12N} & \dots & \dots & a_{N+1N^2} \\ \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots & \dots & \dots & \vdots \\ \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots & \dots & \dots & \vdots \\ a_{N^21} & a_{N^22} & \dots & a_{N^2N} & a_{N^2N+1} & \dots & a_{N^22N} & \dots & \dots & a_{N^2N^2} \end{pmatrix} \quad (52)$$

7. Vypočítá se matice  $\tilde{g}$  podle vztahu (53)

$$\tilde{g} = A(g - \mu_g) \quad (53)$$

8. Získání původního obrazu lze pomocí inverzní KLT (54)

$$g = A^T \tilde{g} - \mu_g \quad (54)$$

Kompresi obrazu se provede tak, že se položí méně významná vlastní čísla autokovarianční matice  $C_{\tilde{g}\tilde{g}}$  rovno nule.

## 5 NÁVRH MODELU

Tato kapitola se bude zabývat již konkrétními daty a využitími v praxi. První část této kapitoly bude věnována teoretickému uvedení do problematiky. Následně budou zpracovány získaná data z veřejně dostupných zdrojů a generovaný signál. Zpracování dat bude provedeno v prostředí MATLAB a MS Excel. Při vytváření řešení problematiky budou dodrženy teoretické zásady, které budou uvedeny v úvodní části této kapitoly. Na data budou aplikovány postupy, které byly nastíněny v předchozích kapitolách teoreticky.

První model se bude zabývat zpracováním ekonomické ČR. Jako sledovaný ekonomický ukazatel bude index spotřebitelských cen (CPI). Data budou naměřena za jednotlivá čtvrtletí od 1. čtvrtletí roku 2000 do 1. čtvrtletí roku 2013. Data budou získána z veřejně dostupné databáze českého statistického úřadu. Cílem této kapitoly bude nalezení vhodného modelu dle Box-Jenkinsovy metodologie.

Další část se bude zabývat zpracováním signálu, který bude vygenerován pomocí funkcí MATLABu, bude u něho zjištěna periodicitu dle postupu uvedeného v teoretické části této práce a dále na tento signál bude aplikováno využití neuronové sítě za účelem zjištění matematického zápisu autokorelační funkce. Tato kapitola bude mít za cíl ukázat zjištění periodicity signálu za využití ACF a naučit neuronovou síť zjistit matematický zápis ACF.

Poslední část této kapitoly bude věnována komprese obrazu za použití Karhunen-Loeveho transformace. Bude zjišťován vliv komprese na kvalitu obrazu, dále bude zjišťováno, jak lze optimalizovat výpočet a poslední model bude ukazovat zpracování barevného obrazu, kde je zapotřebí nejprve obraz rozložit na jednotlivé barevné vrstvy. Cílem této kapitoly je ukázat kompresi obrazu s využitím ACF pro tři různé obrazy a problematiky – vliv na kvalitu obrazu, optimalizace výpočtu a barevný obraz.

### 5.1 Modelování

Model je určitá reprezentace skutečnosti s jistou mírou nepřesnosti. Tudíž je model jistou abstrakcí reality. Vhodným příkladem je mapa, nikdy nezachytí i ty nejjemnější detaily krajiny, terénu, objektu apod. Proces modelování lze dle [34] rozdělit do šesti fází:

1. Formulace problému
2. Základní návrh modelu
3. Implementace modelu

4. Verifikace a validace
5. Simulace a analýza
6. Sumarizace výsledku

### **5.1.1 Formulace problému**

V této fázi je důležité jasně formulovat problém, který se snažíme modelem řešit. Problém je potřeba zachytit, co nejkonkrétněji a pokud možno v číslech, tedy numericky. Musí být určen časový horizont, v němž daný problém studujeme.

### **5.1.2 Základní návrh modelu**

V této fázi se ještě model udržuje co nejjednodušší. Určují se tu prvky, které budou modelovány se zaměřením na hlavní prvky a vztahy mezi nimi. Zjištění prvků s nejdůležitějším vztahem ke studovanému problému.

### **5.1.3 Implementace modelu**

Zde se uskutečňuje výběr vhodného modelovacího přístupu a nástroje, který bude použit. Musí být jasně dané všechny předpoklady. Výstupem této fáze je konkrétní rovnice, tabulky nebo graf, který nějak implementuje daný problém, a pomůžou uskutečnit další kroky.

### **5.1.4 Verifikace a validace**

V této fázi se již ověřuje spustitelný model. Ověřuje se správnost použití modelu, zda zobrazuje chování reálného systému. Ujišťuje se o vztahu mezi realitou a abstraktním návrhem. Dle Pelánka [34] je verifikace ověření, že model dělá to, co by dle předpokladů dělat měl, tedy ověření vztahu mezi abstraktním návrhem modelu a jeho konkrétní implementaci. Jelikož je každý model do jisté míry špatně, zabývá se validace užitečností modelu namísto jeho správnosti.

### **5.1.5 Simulace a analýza**

Cíle analýzy jsou individuální s vazbou na konkrétní problém. Především je třeba za pomoci analýz zjistit role jednotlivých prvků modelu. Které prvky nejvíce ovlivňují chování modelu.

### **5.1.6 Sumarizace výsledku**

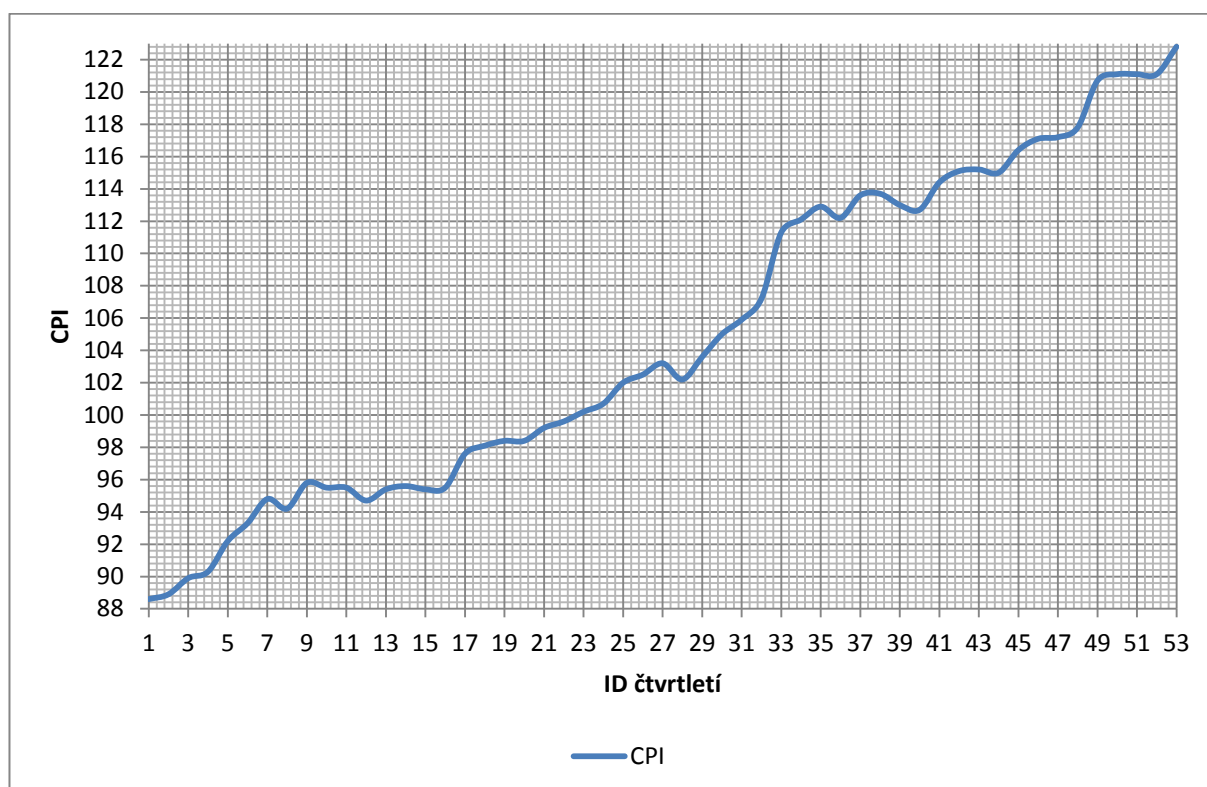
Tato fáze studuje konkrétní výsledky a formulují se tu odpovědi na původní kladené otázky z první fáze. Pokud je potřeba, tak se lze navrátit k předchozím fázím.

## 5.2 Index spotřebitelských cen

Index spotřebitelských cen CPI (angl. Consumer price index) je souhrnný ukazatel růstu cen veškerého zboží a služeb, které domácnosti kupují. Zkoumá se spotřební koš statků a služeb, které domácnosti běžně spotřebovávají. Spotřební koš je tvořen několika sty položek, které vystupují jako reprezentanti v určitém poměru, v jakém se běžně kupuje (např. pečivo se kupuje častěji než televize). Každý reprezentant má v uvedeném koši svou váhu, kterou určuje podíl výdajů na daný statek na celkových výdajích domácnosti. Váhy bývají po určitou dobu fixní, z důvodu náročného statistického zjišťování [35].

V této kapitole bude zpracována ČR představující CPI pro Českou republiku za léta 2000 – 2013. Tato řada je získána z veřejně dostupné databáze Českého statistického úřadu. Zpracovávaná data a zdrojové kódy k tomuto modelu jsou obsažena v příloze A. Cílem této kapitoly bude nalezení vhodného modelu dle Box-Jenkinsovy metodologie.

Data jsou měřena čtvrtletně. Jedná se o 52 pozorování, začínající 1. čtvrtletím roku 2000 a končící 1. čtvrtletím roku 2013. Průběh dat je zobrazen na grafu 1, kde je patrný trend řady, ale sezónní složka není vykazována.

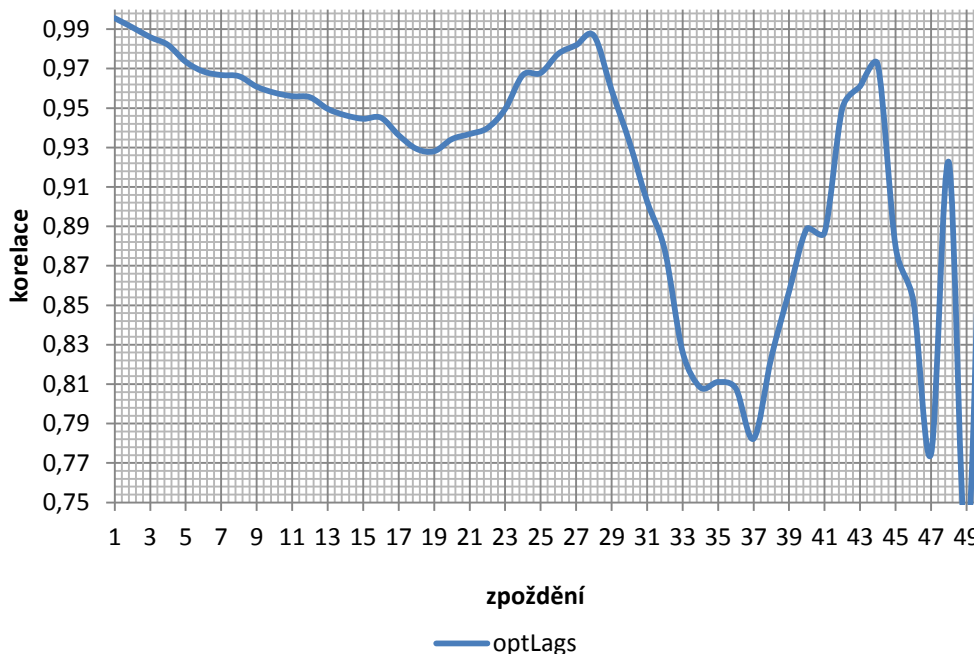


Graf 1: Průběh dat CPI

Zdroj [vlastní]

Již z grafu 1 lze odhadnout, že se zřejmě nebude jednat o stacionární ČŘ, což je ještě nutno ověřit za pomoci autokorelační a parciální autokorelační funkce. Z tohoto důvodu je nutno provést výpočet ACF a PACF, což je v MATLABu provedeno následovně:

1. zjištění optimálního zpoždění (optLags), čímž se zjistí parametr nLags pro funkci autocorr () a parcorr(), což se určí dle mezí odhadnutých podle grafu 2, v tomto případě se jedná o meze <20;30>, tam se nachází hledané maximum (viz kapitola 3.3.3)



**Graf 2:** Zjištění meze optLags

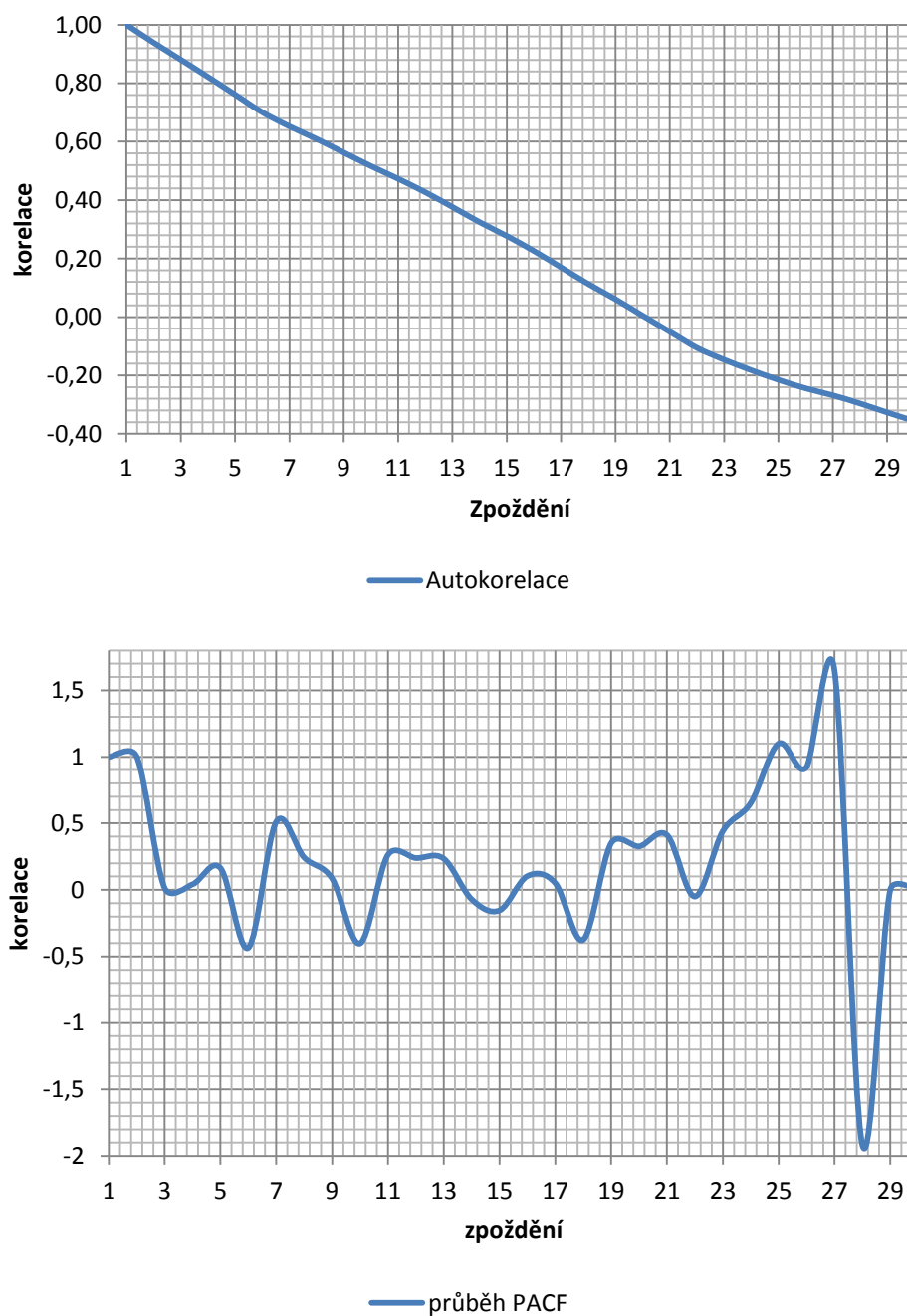
*Zdroj [vlastní]*

2. nyní je určeno nLags dle následujícího kódu:

```
optLags = zeros(n-1,1);
for nLags = 1:1:n-1
    optLags(nLags,1) = corr(CPI(1:end-nLags,2),
    CPI(nLags+1:end,2));
end
plot(optLags);
figure;
[~,nLags] = max(optLags(20:30));
nLags = nLags + 20;
```

3. výsledný parametr nLags je pro tento případ 29, nyní je určen průběh ACF a PACF, což je provedeno dle:

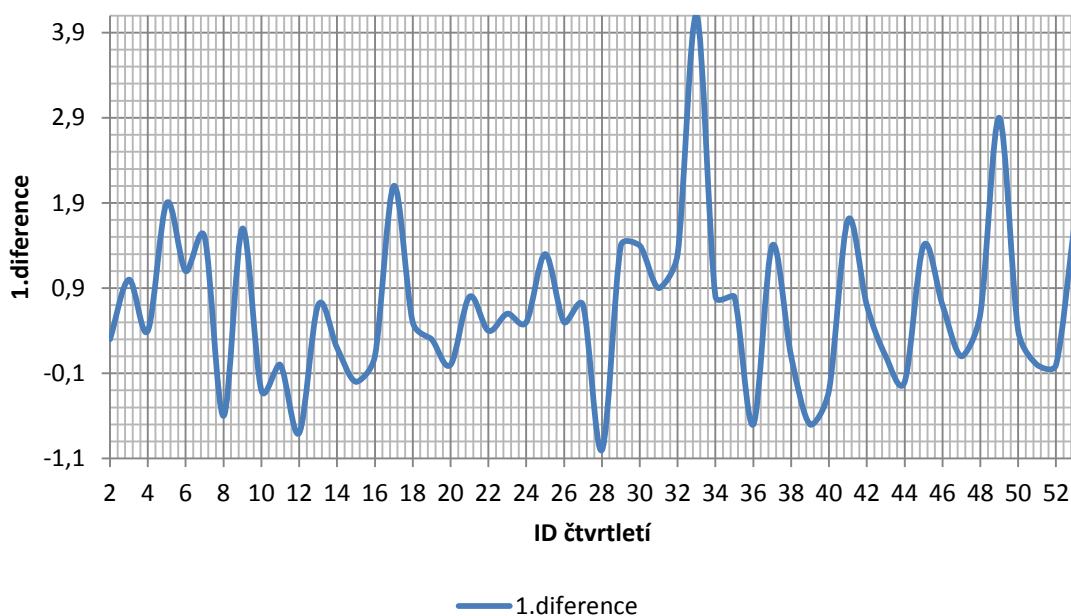
```
[ACF,~,~] = autocorr(CPI(:,2), nLags, 3);  
plot(ACF);  
figure;  
[PACF,~,~] = parcorr(CPI(:,2), nLags, 3);  
plot(PACF);
```



**Graf 3:** Průběh ACF a PACF

*Zdroj [vlastní]*

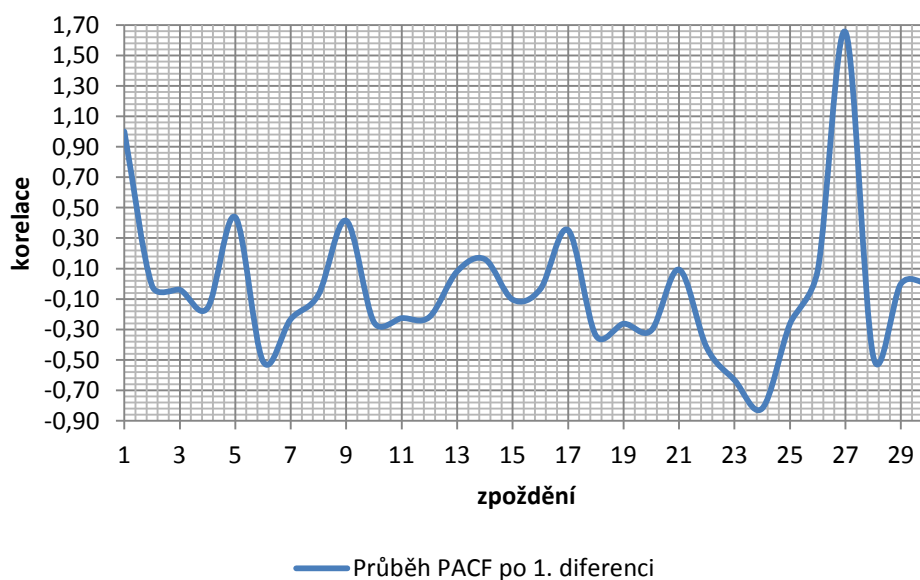
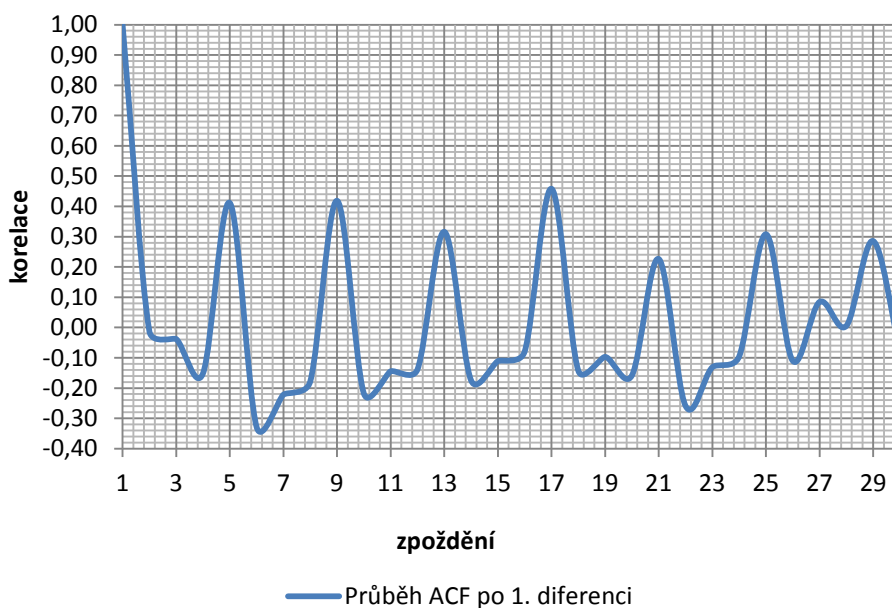
Z grafu 3 je patrné, že autokorelační funkce lineárně klesá, což potvrzuje nestacionaritu ČŘ. Pro odstranění nestacionarity řady je aplikována 1. diference a opět je řada zobrazena graficky (viz graf 4).



**Graf 4:** diferencovaná data

*Zdroj [vlastní]*

Na grafu 4 je již patrné, že je daná řada stacionární, pro ověření je opět zjištěn dle stejného postupu průběh ACF a PACF (viz graf 5).

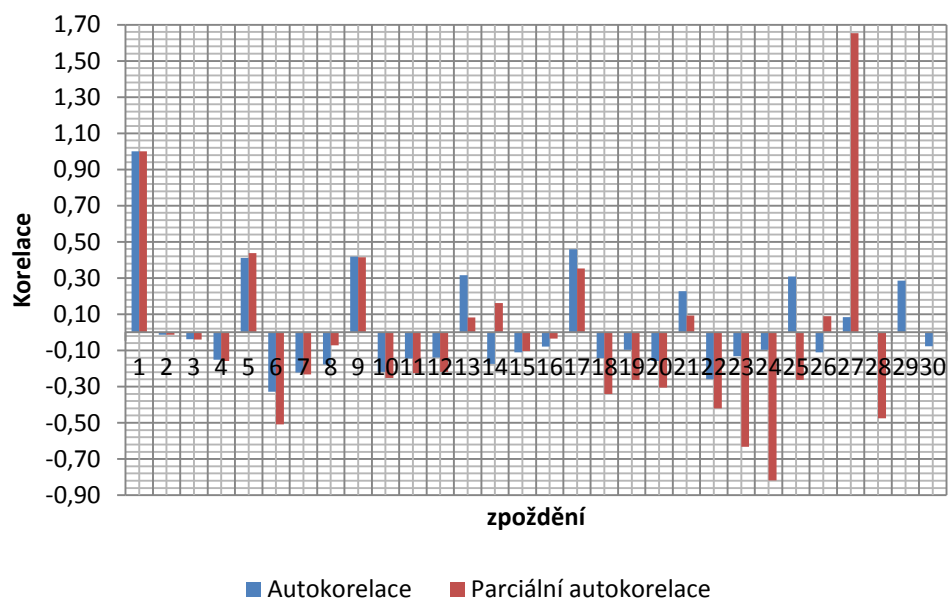


**Graf 5:** ACF a PACF po 1. diferenci

*Zdroj [vlastní]*

Z průběhu ACF a PACF na grafu 5 je čitelné, že 1. diference opravdu odstranila nestacionaritu a řada je nyní stacionární, čímž je splněna hlavní podmínka Box-Jenkinsovy metodologie. V tomto okamžiku již lze přejít k identifikaci modelu.

Pro lepší identifikaci modelu je průběh ACF a PACF po 1. diferenci zobrazen do jednoho korelogramu (viz graf 6).

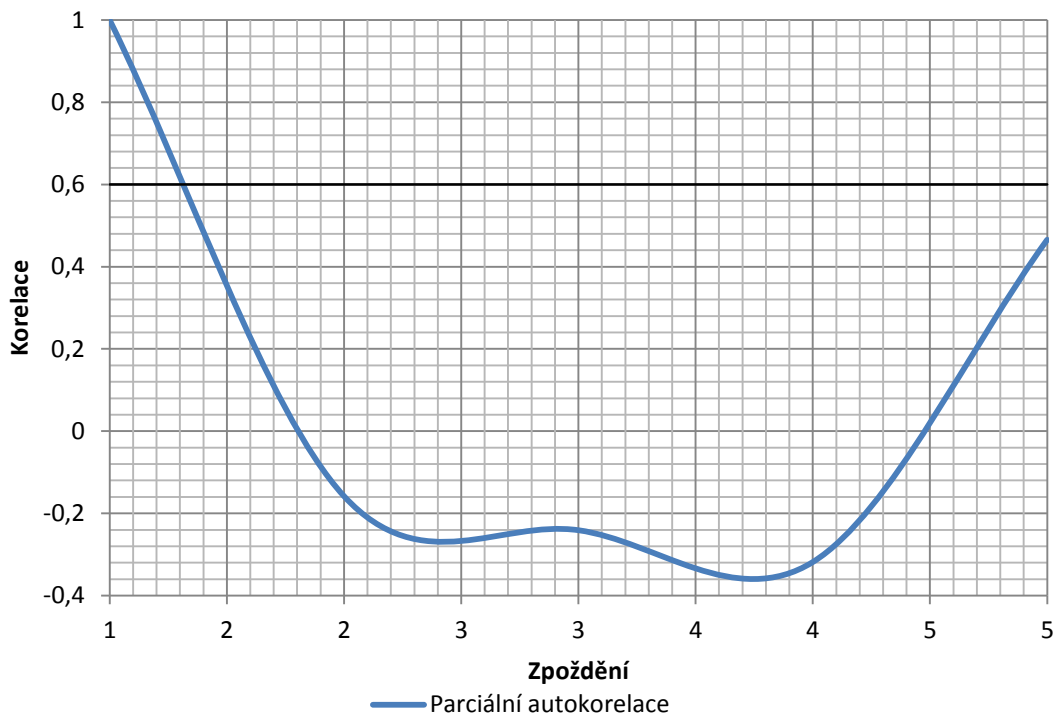
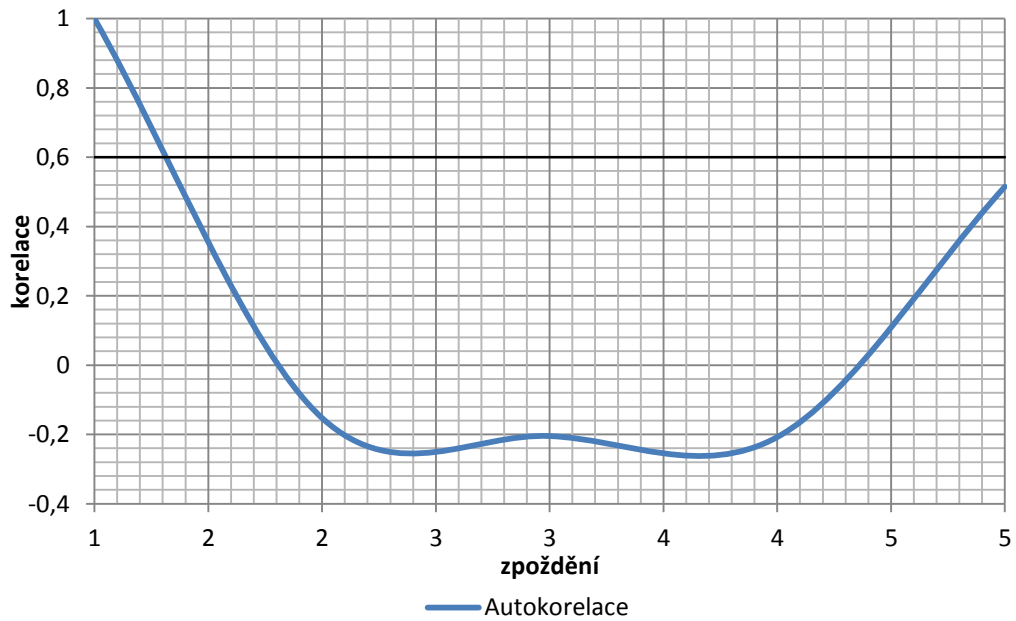


**Graf 6:** Srovnání ACF a PACF

*Zdroj [vlastní]*

Z průběhu ACF a PACF na grafu 6 je patrné, že průběh funkcí je shodný až do zpoždění 3. Na základě subjektivního posuzování a údajů z tabulky 1 lze konstatovat, že toto chování odpovídá autoregresnímu procesu 3. řádu (AR(3)).

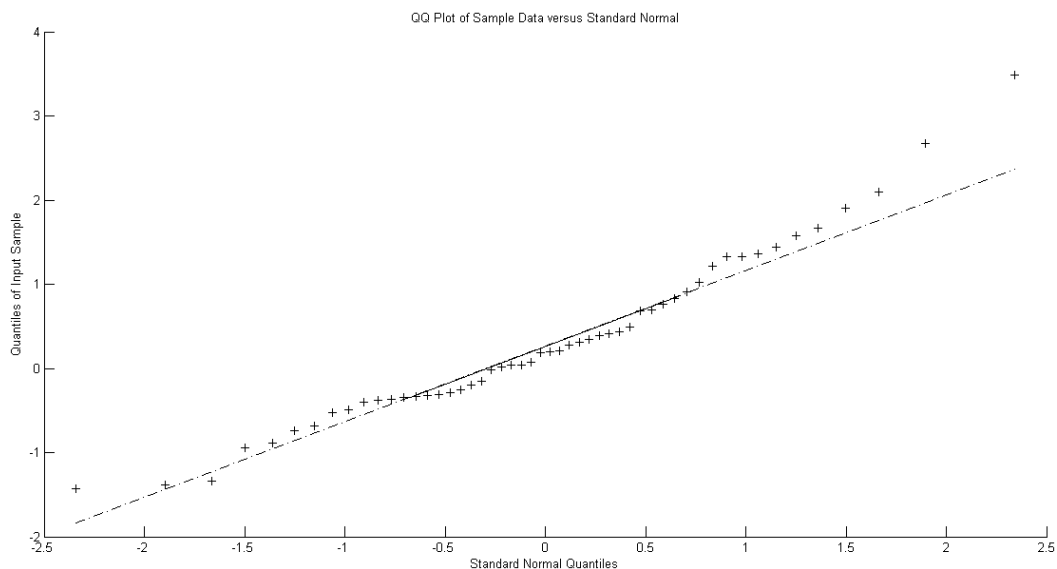
Pro ověření správnosti zvoleného modelu je zapotřebí mít nekorelovaná rezidua s normálním rozložením.



**Graf 7:** Nekorelovanost reziduí

*Zdroj [vlastní]*

Nekorelovanost je zjištěna za pomoci průběhu ACF a PACF reziduí, bere se v úvahu, že data považovaná, jako nekorelovaná, jsou do korelace 0,6. V grafu 7 je vidět, že se opravdu jedná o nekorelovaná rezidua.



**Graf 8:** Q-Q graf reziduí

*Zdroj [vlastní]*

Normalita rozložení reziduí je zjištěna pomocí Q-Q grafu reziduí (viz Graf 8). Graf 8 dokazuje, že je tvrzení o normalitě rozložení reziduí pravdivé a tedy i zvolený model je správný.

Model, rezidua a normalita (Q-Q graf) jsou zjištěny pomocí kódu zapsaného následovně:

```
y = difference1(:,2);
rad = 3;
m = ar(difference1(:,2), rad);
r = resid([y(1:rad);y], m);
r = r(rad+1:end);
qqplot(r);
```

Výsledný model je vygenerován MATLABem následovně:

$$A(q)y(t) = e(t)$$

Vyjde-li se ze vztahu (26) v kapitole 2.3.2, pak jej MATLAB obecně generuje následovně:

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

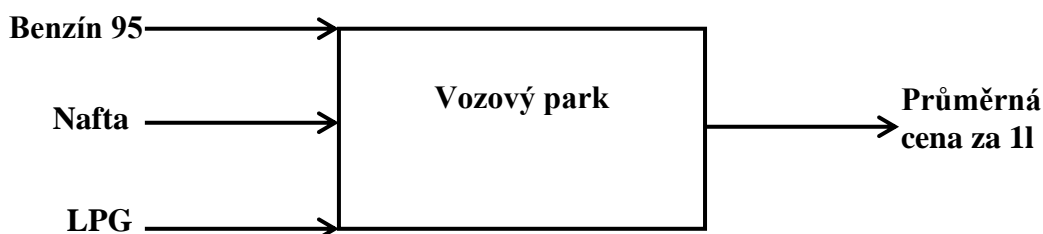
Pro tento případ, jej vygeneroval s konkrétními parametry ve tvaru:

$$A(q) = 1 - 0.2295 q^{-1} - 0.2025 q^{-2} - 0.09273 q^{-3}$$

Cíl zjistit vhodný model dle Box-Jenkinsovy metodologie byl splněn.

### 5.3 Model pro vozový park

Tato kapitola se zabývá případovou studií vycházející z údajů o hypotetické firmě vlastníci vozový park. Firma vlastní a využívá 2 dieslové nákladní vozidla, 8 dieslových dodávek, 7 osobních vozidel na LPG a 6 benzínových osobních vozidel. Situace je zobrazena na obrázku 7. Dále tento model vychází z dat získaných z webového zdroje [36]. Data uvádějí vývoj cen pohonných hmot, které jsou měřeny po týdnech, tudíž jsou pro každý měsíc 4 záznamy. Záznamy jsou pro období 3. 1. 2011 – 22. 7. 2013. Ceny jsou uváděny v eurech za 1litr. Znalosti pro zpracování tohoto modelu byly získány z publikací [37] a [38].

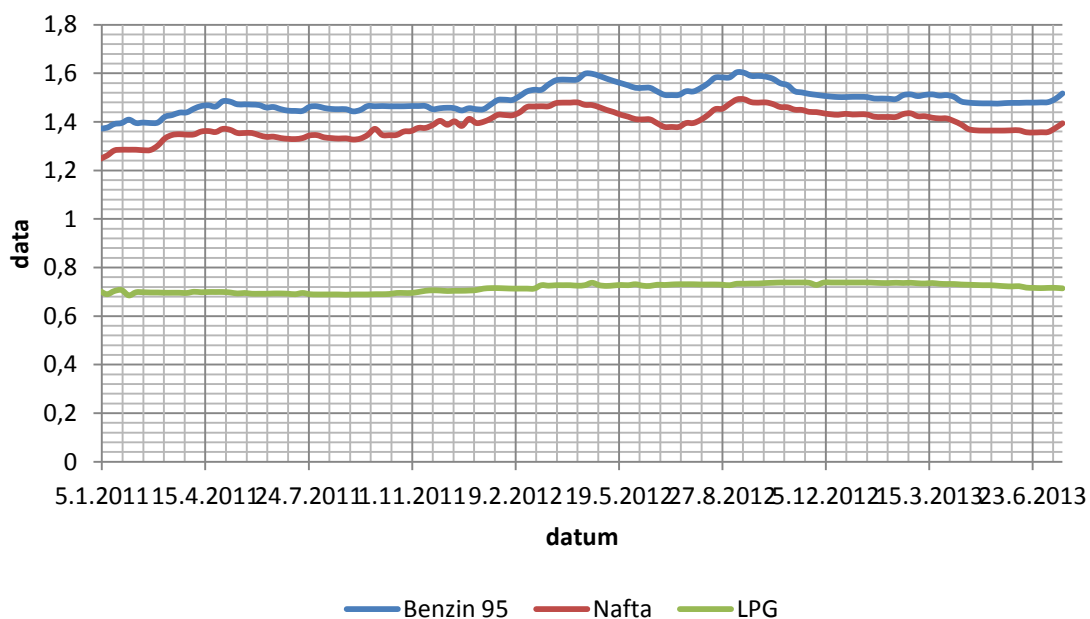


Obrázek 7: Zobrazení modelu

*Zdroj [vlastní]*

Úkolem tohoto modelu je zjištění, kdy se vyplatí nakoupit pohonné hmoty do zásoby.

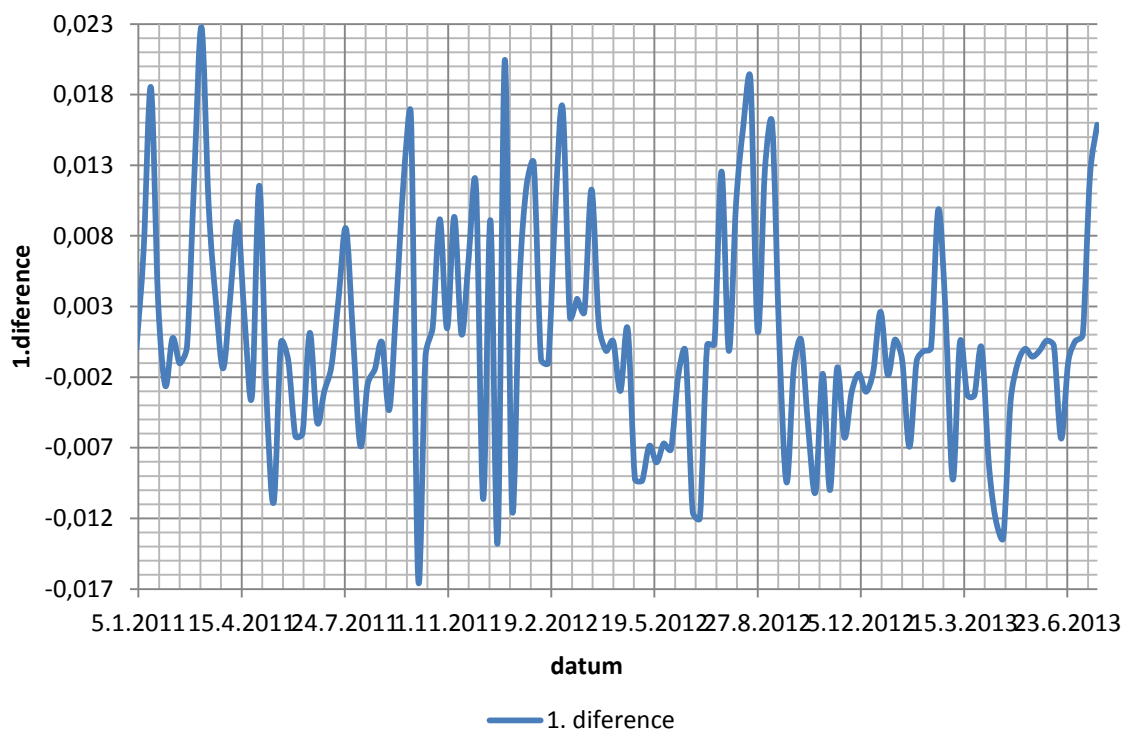
Jelikož se nejedná o stejně hodnocené a využívané suroviny, je nutno určit váhy. Pro určení jednotlivých vah je zapotřebí uvést průměrnou spotřebu vozidel. Nákladní vozidla mají spotřebu  $40l/100km$  nafty, dodávky  $15l/100km$  nafty, osobní vozidla  $9l/100km$  LPG a  $7l/100km$  benzín. Pro naftu se váha určí pomocí součtu součinů počtu vozů a spotřeby, tedy  $2*40+8*15$ , pro LPG je to  $7*9$  a pro benzín  $6*7$ . Z tohoto je možno již určit váhovou funkci, a to prostřednictvím váženého průměru. Váhy a váhová funkce jsou vypočteny v souboru **benzin\_nafta\_lpg.xlsx** a označeny, jako  $w_b$ ,  $w_n$  a  $w_l$  pro váhy a  $y$  pro váhovou funkci (vážený aritmetický průměr).



**Graf 9:** Průběh pohonných hmot

*Zdroj [vlastní]*

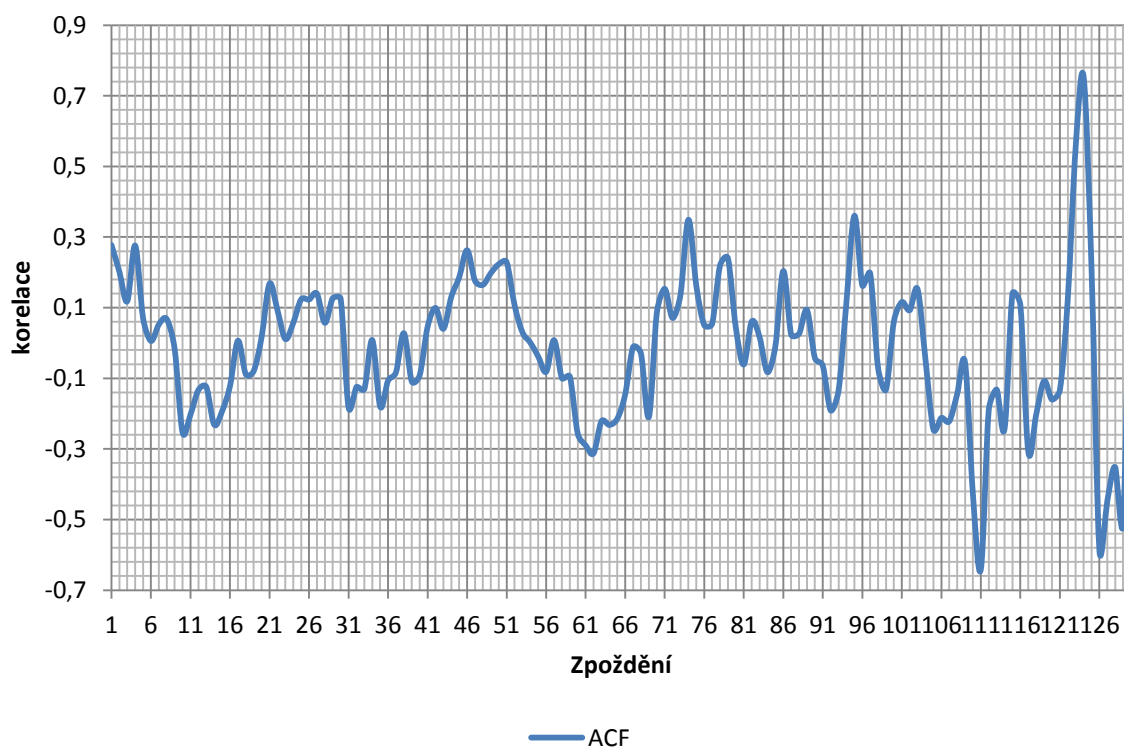
Po zobrazení dat do grafu 9 lze konstatovat, že se jedná o nestacionární data. Jelikož se jedná o nestacionární data je zapotřebí určit první diferenci, aby byla data stacionární a mohlo se pokračovat v analýze s využitím ACF.



**Graf 10:** Zobrazení 1. difference

*Zdroj [vlastní]*

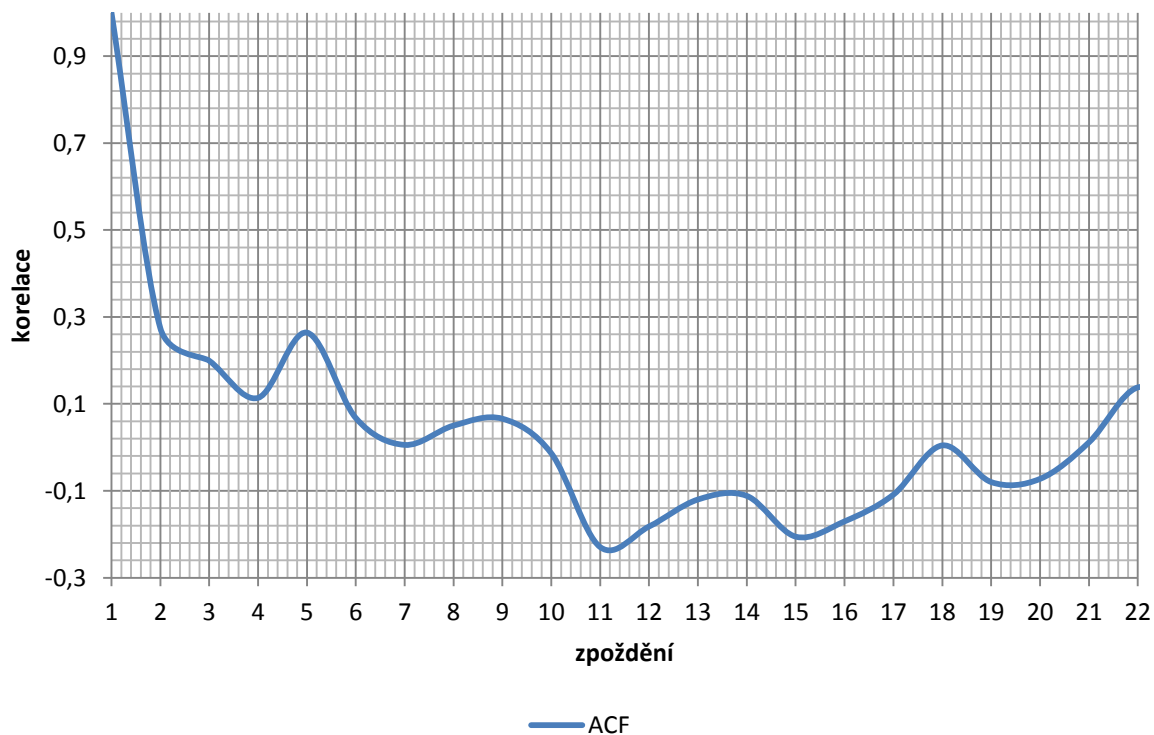
Z grafu 10 je patrné, že se již jedná o stacionární řadu. Nyní lze přistoupit k další analýze a zjistit optimální zpoždění, na jehož základě se určí ACF.



**Graf 11:** Hledání optimálního zpoždění

*Zdroj [vlastní]*

V grafu 11 je vidět, že se optimální zpoždění nachází v intervalu  $\langle 15; 25 \rangle$ . Na základě tohoto intervalu je určen konečný výstup zobrazený v grafu 12.



**Graf 12:** Průběh ACF pohonných hmot

*Zdroj [vlastní]*

Z grafu 12 je čitelné, že nejlevnější litr pohonných hmot pro uvedený vozový park se bude nacházet okolo zpoždění 11, kde je nejznatelnější pokles (údaj ve sloupci ID pro MATLAB na listu „data“ v souboru **benzin\_nafta\_lpg.xlsx**).

**Tabulka 2:** Nalezení nejvýhodnějšího období

<b>Datum</b>	<b>ID pro MATLAB</b>
22.7.2013	1
15.7.2013	2
8.7.2013	3
1.7.2013	4
24.6.2013	5
17.6.2013	6
10.6.2013	7
3.6.2013	8
27.5.2013	9
20.5.2013	10
13.5.2013	11
6.5.2013	12
29.4.2013	13
22.4.2013	14
15.4.2013	15
8.4.2013	16
1.4.2013	17
25.3.2013	18
18.3.2013	19
11.3.2013	20
4.3.2013	21
25.2.2013	22
18.2.2013	23
11.2.2013	24
4.2.2013	25

Při nahlédnutí do zmíněných dat, jejichž ukázka je uvedena v tabulce 2, lze dojít k závěru, že se firmě vyplatí nakoupit pohonné hmoty v polovině měsíce. V grafu 12 je patrné, že poklesy se nacházejí ve zpoždění 2, 7, 11, 15, 19, což potvrzuje výhodnost nákupu v polovině měsíce. Výpočty v MATLABu byly provedeny obdobně, jako v kapitole 5.2, zdrojové kódy se nacházejí v příloze B, soubor s názvem **pohonne\_hmoty.m**.

## **5.4 Generování signálů**

Zdrojové kódy pro tuto kapitolu a následující podkapitolu se nacházejí v příloze C.

### **5.4.1 Zjištění periodicity signálu**

V tomto modelu se bude vycházet ze signálu, které byl vygenerován dle následujícího postupu, pro první signál:

```

function [ y ] = signal1( f0, t )
    y = 0;
    f0 = 2*pi*f0;
    for i = 1:2:10
        y = y + sin(i * f0 * t) / i;
    end
end

```

A pro druhý signál:

```

function [ y ] = signal2( f0, t )
    y = 0;
    f0 = 2*pi*f0;
    for i = 1:4:10
        y = y + sin(i * f0 * t) / i ^ 2 - sin((i + 2) * f0 * t) / (i + 2)
    end
end

```

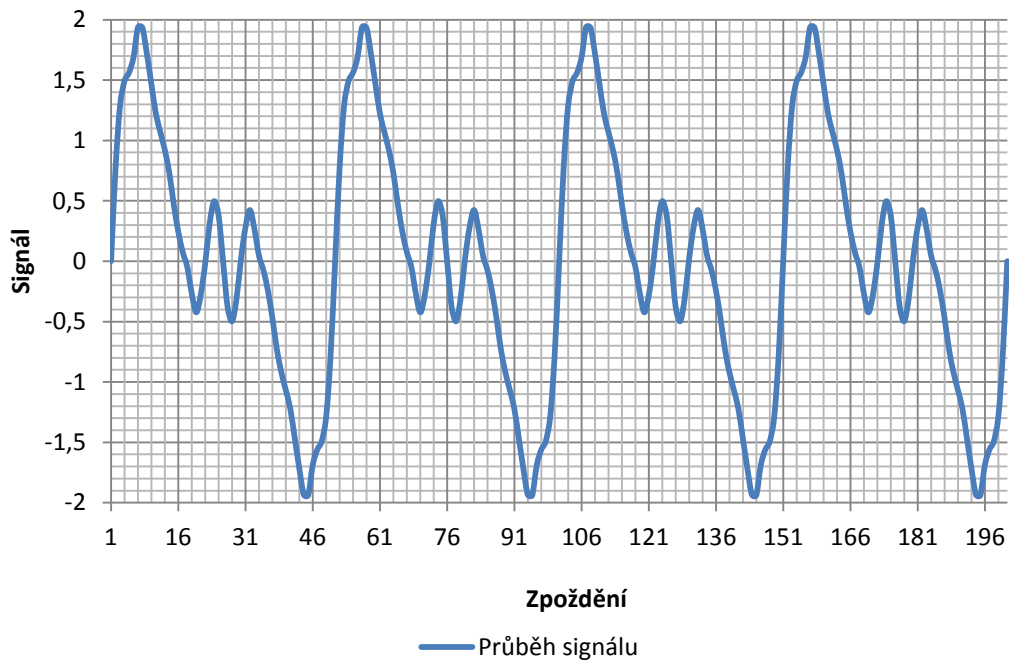
Spojení signálu v jeden signál a vykreslení do grafu:

```

t = 0:0.01:2;
t = t';
n = size(t,1);
s = zeros(n,1);
f0 = 2;
for i = 1:1:n
    s(i,1) = signal1(f0, t(i,1)) + signal2(f0*2, t(i,1));
end
plot(t',s);
figure;

```

Cílem analýzy tohoto signálu bude zjištění periodicity na základě autokorelační funkce. Postup se odráží od teoretické části kapitoly 3.3.1.

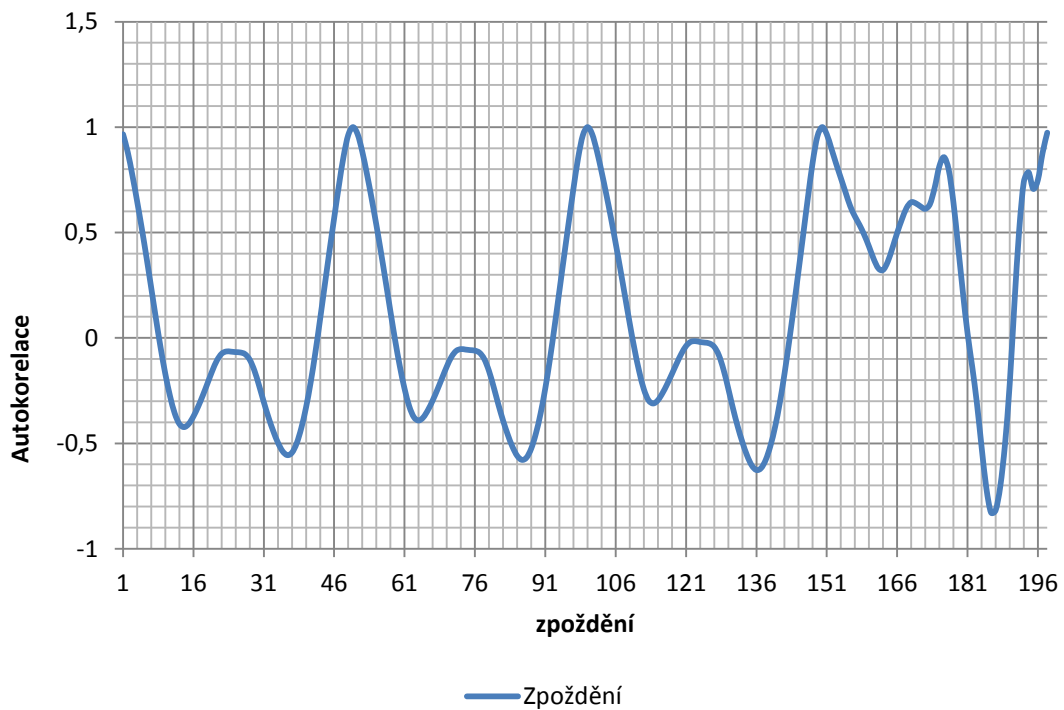


**Graf 13:** Průběh generovaného signálu

*Zdroj [vlastní]*

V grafu 13 je zobrazen průběh vygenerované funkce dle uvedeného kódu. Pro další práci je potřeba určit optimální délku zpoždění (optLags), výchozím pro určení mezí je vykreslení autokorelace do grafu a nalezení maximálního vrcholku hlavního laloku, což lze v prostředí MATLAB provést dle zdrojového kódu:

```
optLags = zeros(n-1,1);
for nLags = 1:1:n-1 %nLags musí být vyšší než řád vyhlazování
    optLags(nLags,1) = corr(s(1:end-nLags,1), s(nLags+1:end,1));
end
plot(optLags);
figure;
```



**Graf 14:** Určení mezí optLags

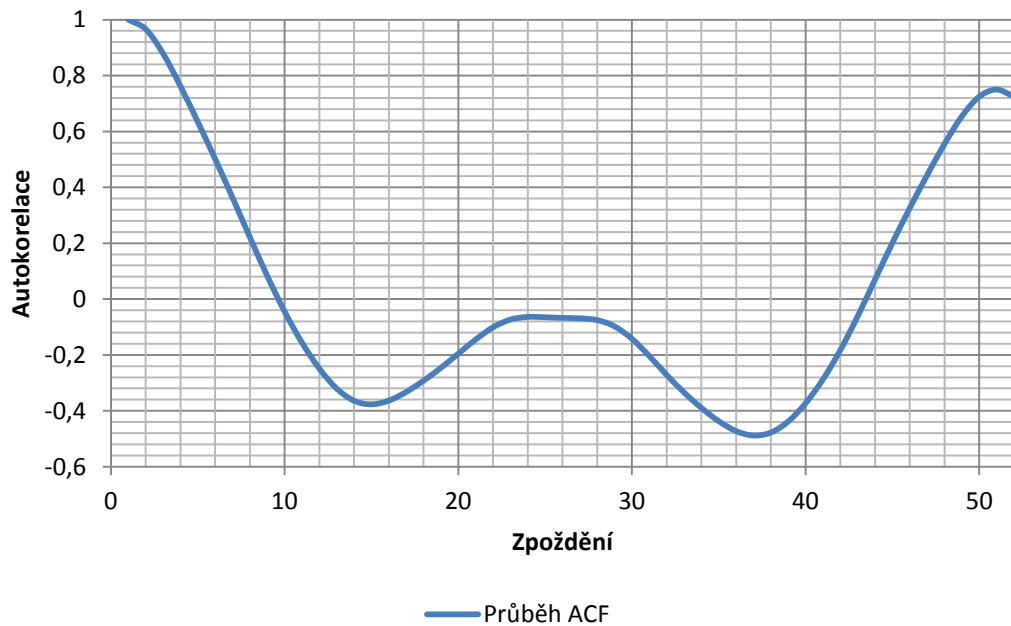
*Zdroj [vlastní]*

Z grafu 14 je patrné, že hledané maximum se bude nacházet v intervalu  $\langle 20;80 \rangle$ , takže hledané zpoždění (nLags), které je zapotřebí k výpočtu autokorelační funkce bude zjištěno na základě tohoto zápisu:

```
[~,nLags] = max(optLags(20:80)); % rozsah byl určen na základě grafu
optLags
nLags = nLags + 20;
```

Výstupním optimem je pro tento signál 52, nyní už tedy lze vypočítat autokorelační funkci a zobrazit její průběh do grafu. Výpočet ACF je proveden dle:

```
[ACF,Lags,Bounds] = autocorr(s(:), nLags, 3);
plot(ACF);
figure;
```



**Graf 15:** Průběh ACF pro určení periodicity signálu

*Zdroj [vlastní]*

Graf 15 ukazuje periodicitu funkce ACF, která je stejná, jako u původního signálu, je vidět, že periodičita nastává ve zpoždění 50, z čehož vyplývá, že délka periodicity je 2, vychází se z nadefinovaných časových posunů, které činí 0,01 ( $=1/(50*0,01)$ ), pro kontrolu musí vyjít periodičita shodná, jako bylo zadáno na počátku generování signálu, což je v tomto případě opravdu 2 Hz.

#### 5.4.2 Odstranění šumu z generovaného signálu

V této kapitole bude ukázáno využití ACF pro detekci šumu v signálu. Pro tento model bude nejprve vygenerován signál s určitou náhodnou složkou, která způsobí jeho zašumění. Tento šum bude následně od signálu oddělen pomocí ACF.

Pro tento model byl vygenerován signál dle uvedeného postupu, pro první signál:

```
function [ y ] = signal1( f0, t )
    y = 0;
    f0 = 2*pi*f0;
    for i = 1:2:10
        y = y + sin(i * f0 * t) / i;
    end
end
```

A pro druhý signál:

```

function [ y ] = signal2( f0, t )
y = 0;
f0 = 2*pi*f0;
for i = 1:4:10
    y = y + sin(i * f0 * t) / i ^ 2 - sin((i + 2) * f0 * t) / (i + 2) ^2;
end
end

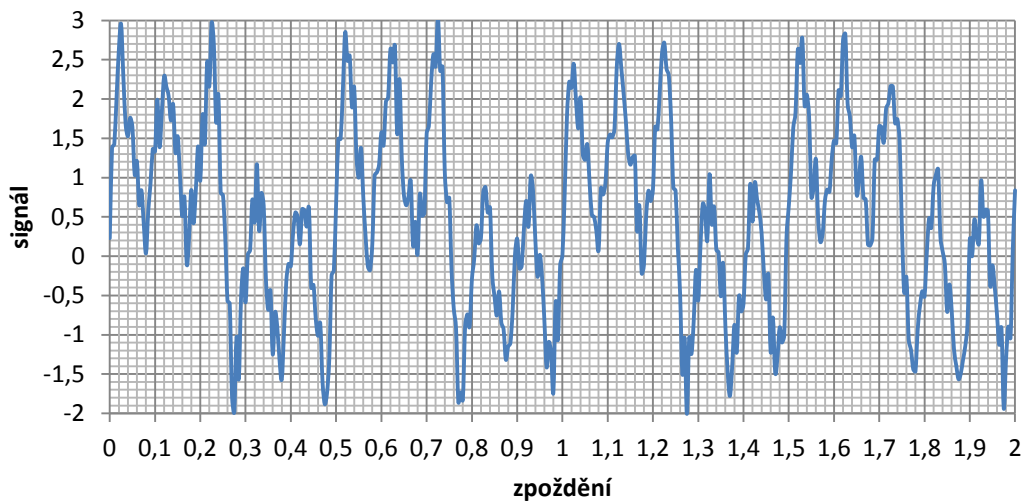
```

Spojení signálu v jeden signál a vykreslení do grafu, na rozdíl od předchozích modelů, je tento signál účelně zašuměn pomocí funkce **rand()**:

```

t = 0:0.005:2;
t = t';
n = size(t,1);
s = zeros(n,1);
f0 = 2;
for i = 1:1:n
s(i,1) = signal1(f0, t(i,1)) + signal2(f0*5, t(i,1))+rand(1);
end
plot(t',s);
figure;

```

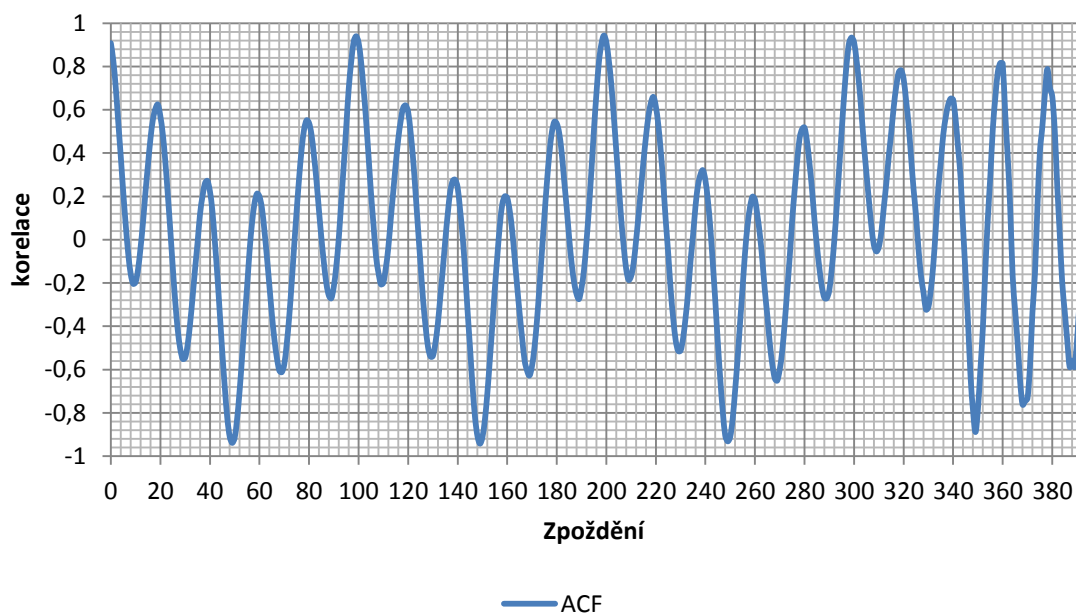


— Průběh zašuměného signálu

**Graf 16:** Vygenerovaný zašuměný signál

*Zdroj [vlastní]*

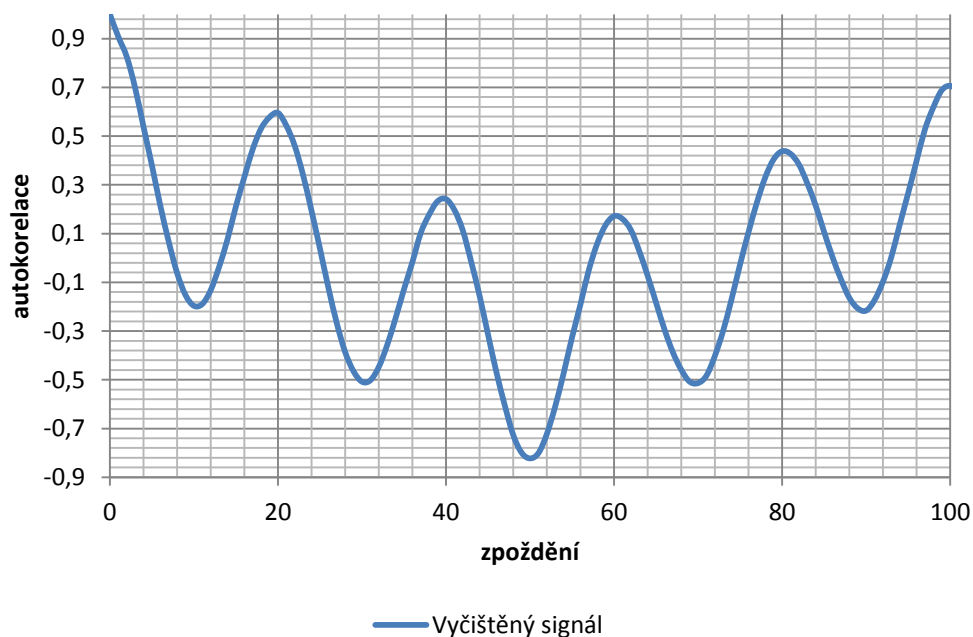
Graf 16 zobrazuje vygenerovaný zašuměný signál, který byl generován MATLABem dle výše uvedeného postupu.



**Graf 17:** Zjištění parametru nLags

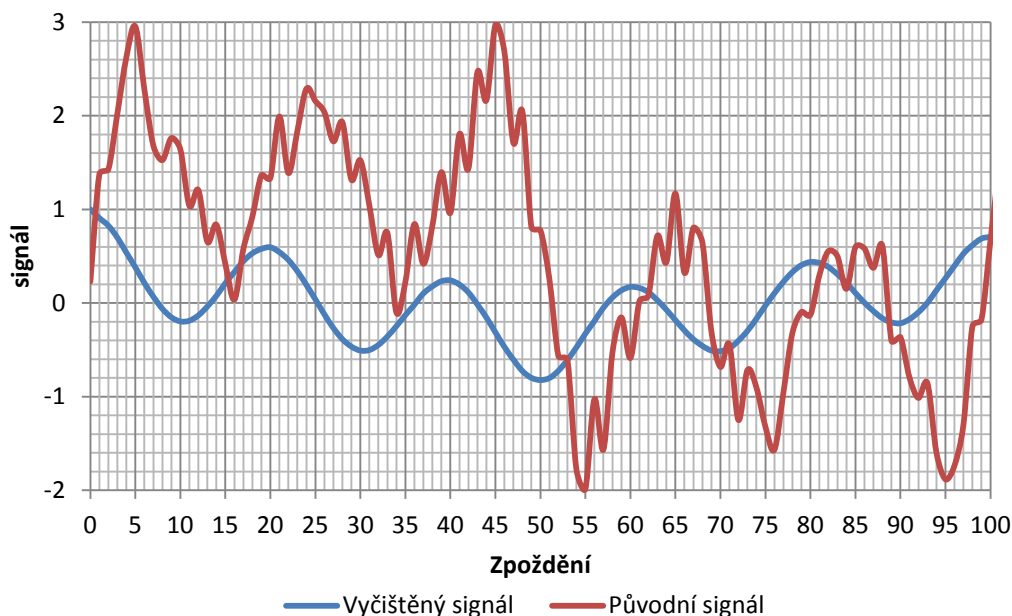
*Zdroj [vlastní]*

Graf 17 slouží ke zjištění parametru nLags pro následné zjištění ACF, která bude pro tento případ zobrazovat vyčištěný signál o šum. Z grafu 14 je patrné, že se optimální zpoždění bude nacházet někde v mezích <75;125>.



**Graf 18:** Vyčištěný signál

V grafu 18 je zobrazen průběh získané ACF, který je v tomto modelu hledaný vyčištěný signál o šum.



**Graf 19:** Srovnání zašuměného a vyčištěného signálu

*Zdroj [vlastní]*

Pokud se porovná vyčištěný a původní signál, tak je patrné, že poklesy a výkyvy vlivem zašumění se opravdu vyhladily a odstranily a zbyl signál bez šumu, což lze zjistit pouhým okem v grafu 19.

V závěru lze tedy konstatovat, že úloha odstranění šumu ze signálu byla učiněna správně a cíl tohoto modelu byl splněn.

## 5.5 Karhunen-Loeveho transformace

V této kapitole je ukázána aplikace KLT na dva různé obrazy. Jedná se o logo UPce a snímek z vesmíru. Logo Upce je úmyslně zmenšeno, a to z důvodu rychlejšího zpracování, jak již bylo uvedeno v kapitole 4.2, tak KLT je velice výpočtově náročná, tedy co se paměti a času týká. Snímek z vesmíru byl rozdělen na 6 částí o velikosti 50x50 pixelů. Po provedení výpočtových operací byly snímky opět spojeny v jeden o velikosti 150x100 pixelů. S KLT se nejčastěji pracuje při kompresi obrazu, při transformaci je využíváno ACF, což je uvedeno v postupu KLT v kapitole 4.2.

Jako výchozí zdrojové kódy pro níže uvedené podkapitoly jsou **autocorr.m**, **m2v.m**, **v2m.m**, **rotc.m** a **KLtrans.m** nacházející se v příloze D.

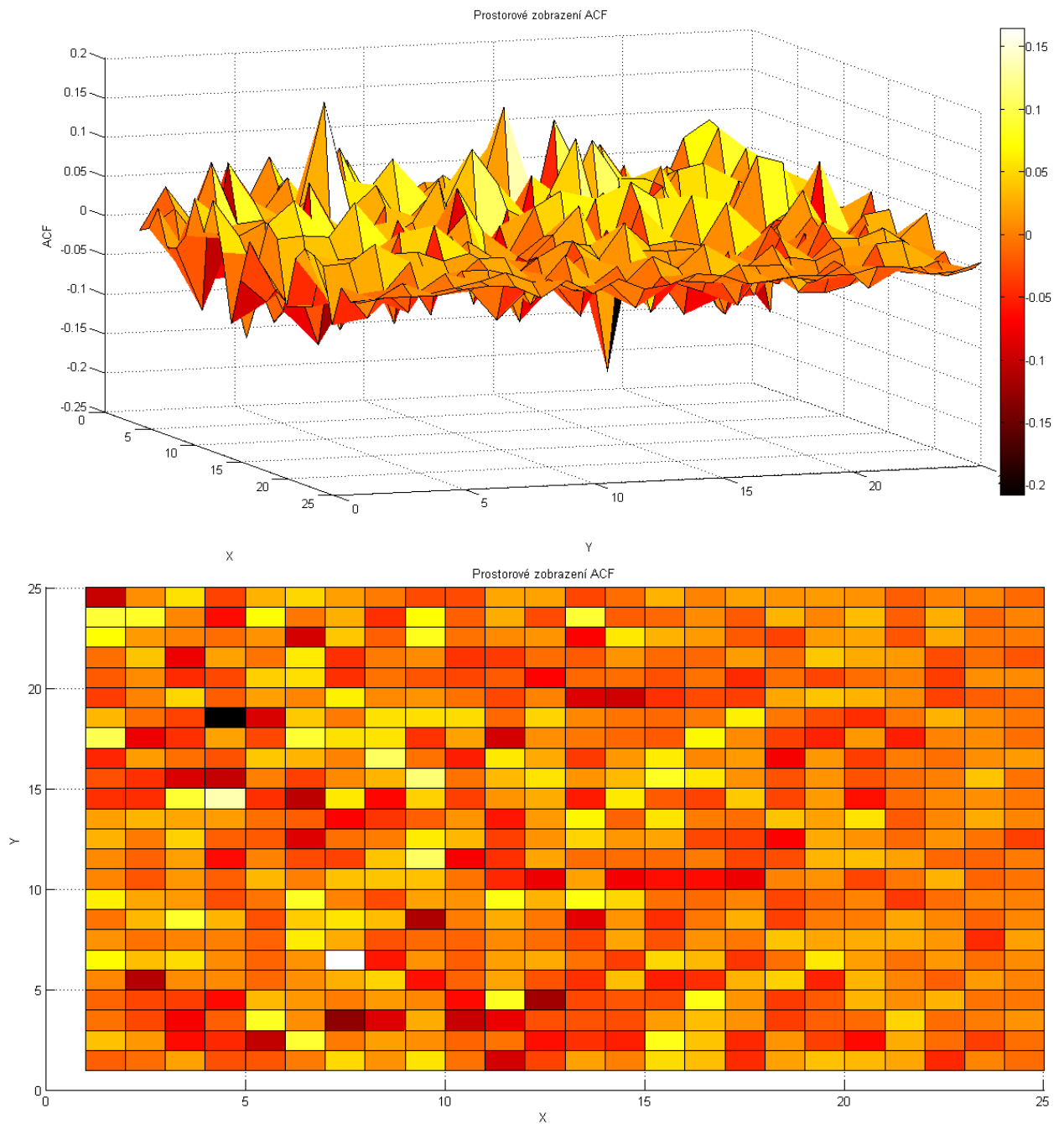
### 5.5.1 Vliv komprese na kvalitu obrazu

Zde bude ukázán vliv komprese na kvalitu obrazu. Platí tu, že čím větší komprese, tím nižší kvalita a naopak. S kompresí od 60-ti do 80-ti % lze pozorovat subjektivní zlepšení zašuměného obrazu. Komprimovaný obraz má velikost 25x25 pixelů.

Pro tento příklad je pro znázornění vykresleno prostorové vyjádření autokorelační funkce matice obrazu. Vykreslení grafu je provedeno dle tohoto zápisu:

```
surf(Macf);  
colormap hot;  
title('Prostorové zobrazení ACF');  
xlabel('X'); %X-ové souřadnice obrazu  
ylabel('Y'); %Y-ové souřadnice obrazu  
zlabel('ACF');
```

Grafický výstup MATLABu a natočený graf, tak aby bylo čitelné maxima (bílá), minima (černá) a nezávislé (oranžová) ACF se nachází na grafu 20.

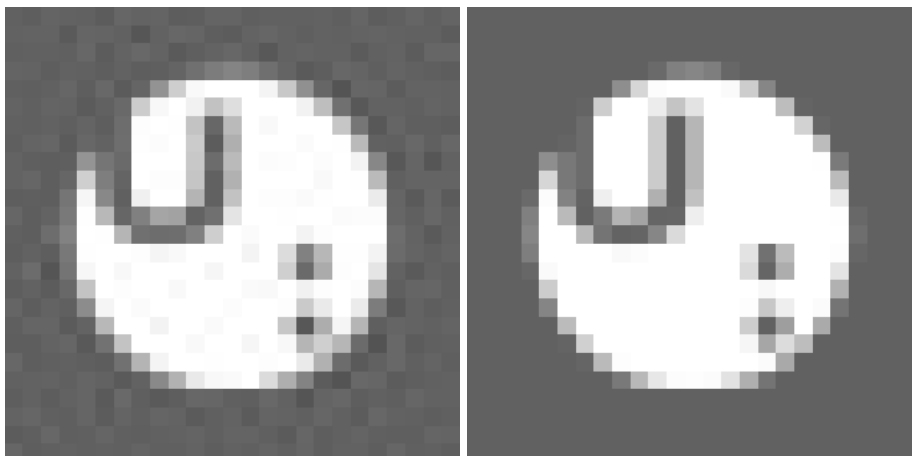


**Graf 20:** Prostorové zobrazení ACF matice obrazu

*Zdroj [vlastní]*

Následující ukázky komprimovaných obrázků jsou rozděleny na pravou a levou část. Originál komprimovaného obrázku se vždy nachází vpravo a komprimovaný obraz za zadaného kompresního poměru je vlevo.

### **Kompresní poměr 40 %**

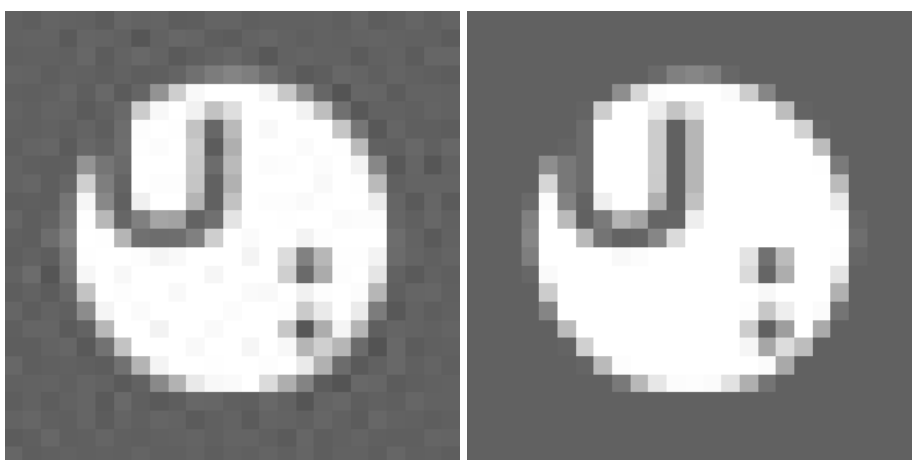


**Obrázek 8:** Srovnání komprese 40% a originálu

*Zdroj [vlastní dle upce.cz]*

Subjektivní vizuální vjem z komprimovaného obrazu, uvedeného v levé části Obrázek 8, je lepší, než bylo očekáváno. Horší výsledek byl očekáván z důvodu ztráty 40 % informací oproti originálu.

### **Kompresní poměr 44 %**



**Obrázek 9:** Srovnání komprese 44% a originálu

*Zdroj [vlastní dle upce.cz]*

Nyní byla porovnávána 44 %-tní rozdílnost mezi předchozím kompresním poměrem (viz obrázek 8) a novým poměrem (viz obrázek 9). Změna byla tak malá, že pouhým okem není zřetelná, ale po bližším náhledu lze spatřit jistou odlišnost ve vykreslení, takže tento rozdíl odpovídá očekávání. Logo je i v tomto případě čitelné.

### **Kompresní poměr 48 %**

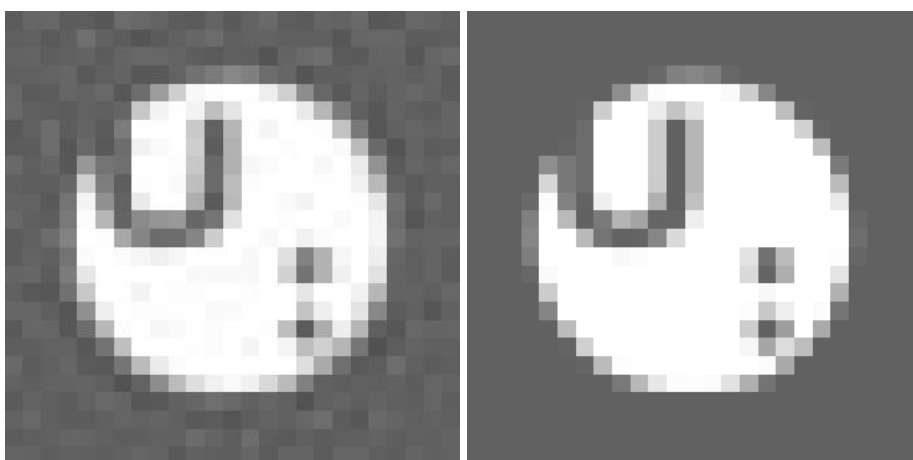


**Obrázek 10:** Srovnání komprese 48% a originálu

*Zdroj [vlastní dle upce.cz]*

Opět byla změna tak malá, že pouhým okem není zřetelná (viz Obrázek 10), ale po bližším náhledu lze spatřit jistou odlišnost ve vykreslení, takže tento rozdíl odpovídá očekávání. Logo je i v tomto případě dobře čitelné.

### **Kompresní poměr 52 %**



**Obrázek 11:** Srovnání komprese 52 % a originálu

*Zdroj [vlastní dle upce.cz]*

Změna oproti předchozímu obrazu je opět minimální, tak jak bylo očekáváno (viz Obrázek 11). Pokud se komprimovaný obraz porovná s první kompresí, pak je rozdíl i pouhým okem znatelný. Logo je stále dobře čitelné.

### **Kompresní poměr 56 %**

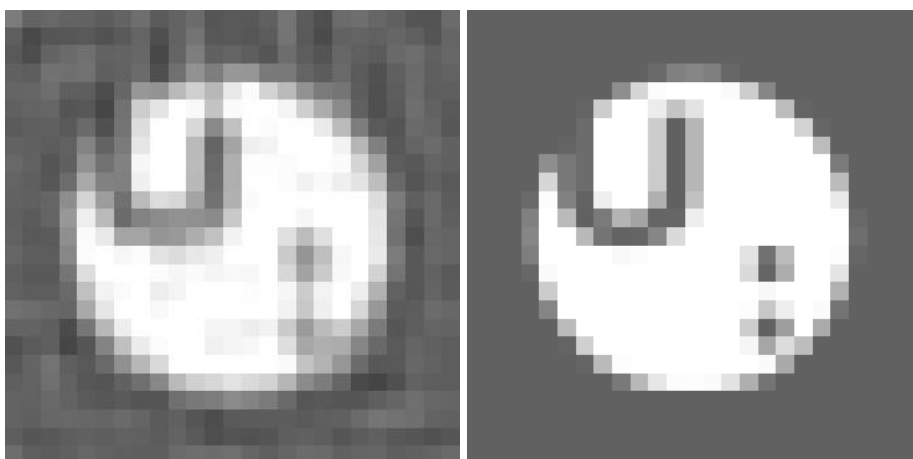


**Obrázek 12:** Srovnání komprese 56 % a originálu

*Zdroj [vlastní dle upce.cz]*

Nyní již lze na obrázku 12 zpozorovat, že dochází ke světlejšímu vykreslení tmavých kruhovitých symbolů na světlém pozadí, ale i přes toto je logo čitelné.

### **Kompresní poměr 80 %**

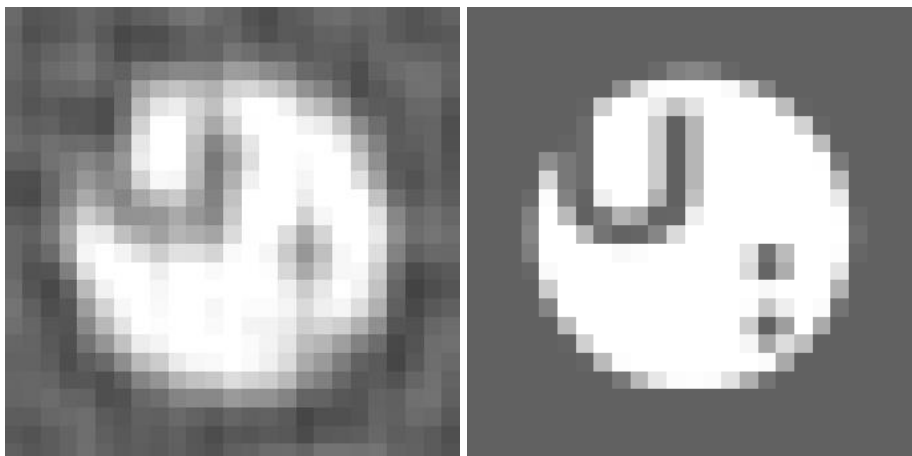


**Obrázek 13:** Srovnání komprese 80 % a originálu

*Zdroj [vlastní dle upce.cz]*

Nyní je rozdíl citelnější (viz obrázek 13), jelikož je tu rozdíl 24 %-tní. Logo je sice čitelné, ale už nejsou jednotlivé útvary, tak dobře vykresleny, aby byli oku lichotivé.

### Kompresní poměr 88 %



**Obrázek 14:** Srovnání komprese 88 % a originálu

*Zdroj [vlastní dle upce.cz]*

Tento poměr je již nevhodný, protože dochází k daleko větší deformaci (viz Obrázek 14), než tomu tak bylo u kompresního poměru 80%.

Závěrem lze říci, že tvrzení, které bylo na začátku uvedeno, je pravdivé, tedy že lze s kompresí od 60-ti do 80-ti % pozorovat subjektivní zlepšení zašuměného obrazu.

#### **5.5.1 Optimalizace procesu zpracování obrazu**

Z důvodu vysoké výpočetní náročnosti se bude tato kapitola zabývat optimalizací tohoto výpočtu, tak aby došlo k rychlejšímu zpracování obrazu. Tento proces bude ukázán na snímku z vesmíru o velikosti 180x120 pixelů. Tento snímek je uveden na obrázku 15.



**Obrázek 15:** Vstupní obraz - vesmir.jpg

*Zdroj [vlastní]*

První fází bylo převedení obrazu na černobílý, což bylo provedeno v rámci zpracování v MATLABu. Tento proces je vidět v následující ukázce zdrojového kódu:

```
obraz = imread('vesmir.jpg'); % načtení obrázku jako matice
obraz = im2double(obraz); % konverze datového typu na double
% převod na stupně šedi
obraz_sedy = (obraz(:,:,1) + obraz(:,:,2) + obraz(:,:,3)) / 3;
kvalita = 0.15; % kvalita obrazu
fragmentMN = 15; % velikost čtvercových fragmentů
```

Nejprve byl obraz rozložen na čtvercové fragmenty o velikosti 40 na 40 obrazových bodů. Toto rozložení však vedlo k velkému počtu operací během zpracování obrazu. Z tohoto důvodu byly fragmenty zmenšeny na velikost 15 na 15 obrazových bodů.

Postup operace je vidět v následující ukázce zdrojového kódu, nejprve operace rozložení na fragmenty a poté spojení v jeden obraz:

```
% postupné rozřezávání a spojování obrazu
for x = 1:fragmentMN+1:M
    nextX = x + fragmentMN;
    for y = 1:fragmentMN+1:N
        nextY = y + fragmentMN;
        if nextX <= M && nextY <= N
            fragment = obraz_sedy_rozsireny(x:nextX,y:nextY); % odříznutí
            [~, fragment] = KLtrans(fragment, kvalita); % komprese fragmentu
            obraz_komprim_rozsireny(x:nextX,y:nextY) = fragment; % připojení
        end;
    end;
end;
```

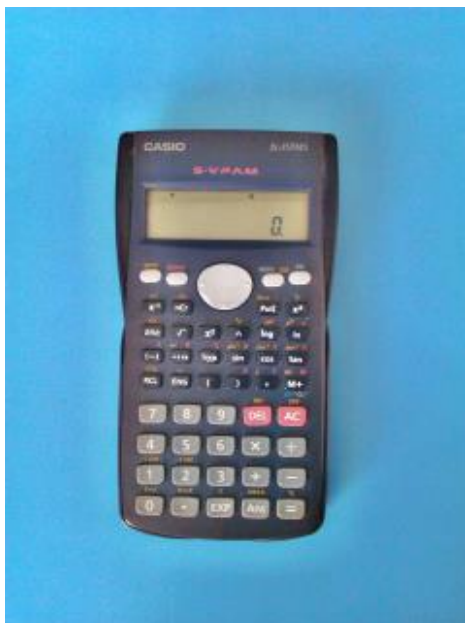
V průběhu komprese je využita autokorelační funkce ke zjištění závislosti každého jednotlivého obrazového bodu na ostatních obrazových bodech. Význam bodu pro celistvost obrazu je dán mírou jeho závislosti na okolních obrazových bodech. Komprese je založena na zanedbávání bodů s nízkou významností pro celistvost obrazu.

Operace nad fragmenty jsou nezávislé, proto je lze vykonávat paralelně. Lze tu využít výhod grafických karet a vícejádrových procesorů.

Změna čtverců ze 40-ti obrazových bodů na 15 obrazových bodů vedla k urychlení výpočtu o 96%, a to bez ztráty na kvalitě. Tedy cíl optimalizace výpočtu byl nad očekávání splněn.

### 5.5.2 Zpracování barevného obrazu

Pro další analýzu bude použit barevný obraz kalkulačky, na kterém bude opět aplikována Karhunen-Loeveho transformace za účelem komprese obrazu. Obraz kalkulačky je uveden na obrázku 16.



**Obrázek 16:** Původní obraz kalkulačky

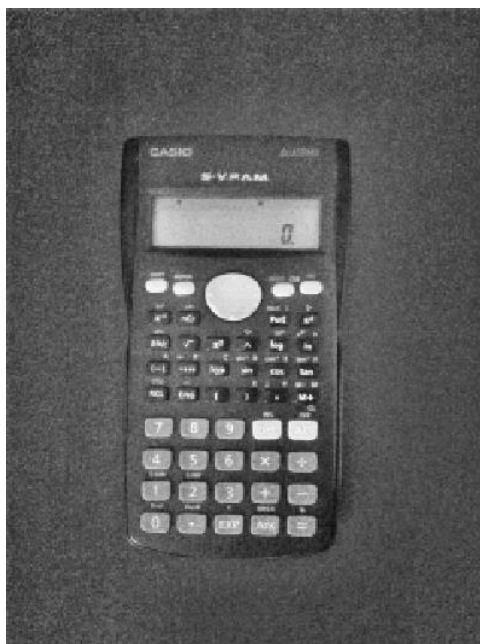
*Zdroj [vlastní]*

Při zpracování barevného obrazu je ovšem zapotřebí rozložit daný obraz na jednotlivé barevné vrstvy (podobrazy) modelu RGB. Tyto vrstvy jsou převedeny na matice, které jsou dále převedeny na vektory, jež se zpracovávají pomocí výše vyjmenovaných operací. Po dokončení operací jsou tyto jednotlivé vektory opět převedeny na matice a spojeny v jeden obraz.

Zápis pro rozdělení obrazu na jednotlivé vrstvy RGB a rozšíření o fragmenty, tak aby byli zkomprimováni i okrajové pixely, je následující:

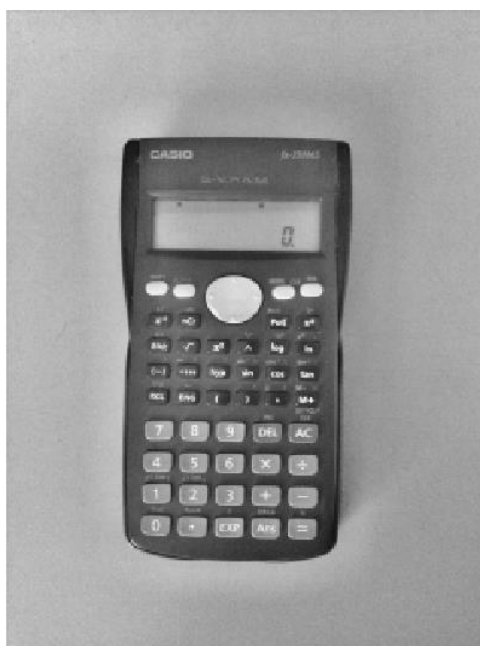
```
[M N D] = size(obraz);  
obraz_komprim = zeros(M, N, D);  
for d = 1:1:D  
    podobraz = obraz(:,:,d);  
    imshow(podobraz, 'InitialMagnification','fit');  
    figure;  
    % rozšíření obrazu  
    M = M + fragmentMN;  
    N = N + fragmentMN;  
    podobraz_komprim_rozsireny = zeros(M, N);  
    podobraz_rozsireny = zeros(M, N);  
    podobraz_rozsireny(1:size(podobraz,1),1:size(podobraz,2)) = podobraz;
```

Výsledný barevný podobraz vrstvy „red“ je uveden na obrázku 17, podobraz „green“ na obrázku 18 a podobraz „blue“ je uveden na obrázku 19.



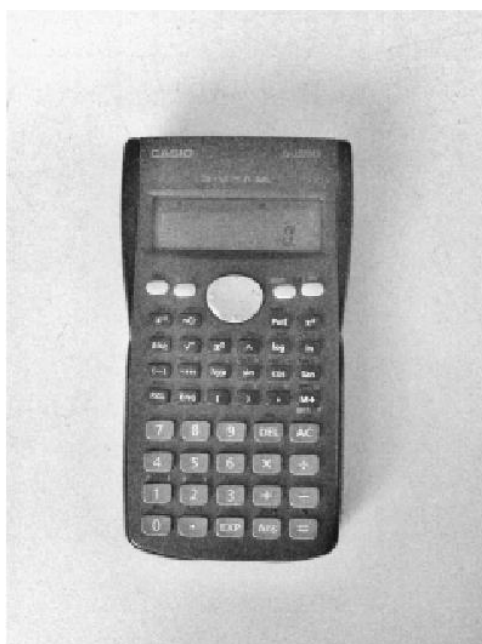
**Obrázek 17:** Podobraz "red" – bez komprese

*Zdroj [vlastní]*



**Obrázek 18:** Podobraz "green" - bez komprese

*Zdroj [vlastní]*

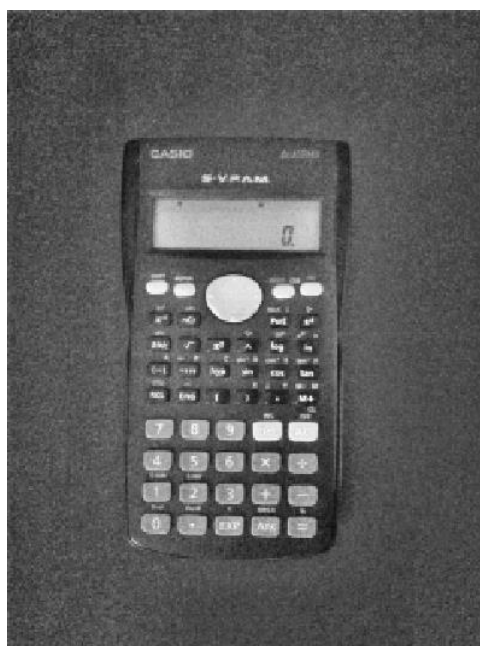


**Obrázek 19:** Podobraz "blue" - bez komprese

*Zdroj [vlastní]*

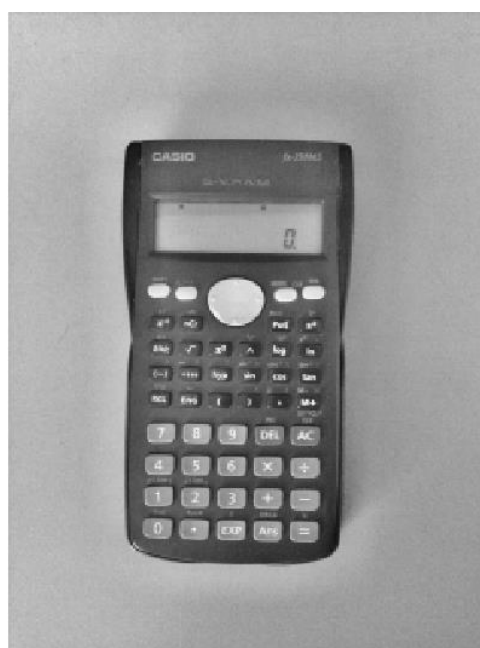
Nyní už je možné provést komprese obrazu, při níž dojde ke kompresi jednotlivých podobrazů a následnému spojení v jeden obraz, tak aby byl opět získán barevný obraz.

Obrázek 20 zobrazuje zkomprimovanou vrstvu „red“, obrázek 21 vrstvu „green“ a obrázek 22 zobrazuje zkomprimovanou vrstvu „blue“. Na obrázku 23 je uvedený koncový výstup komprese obrazu.



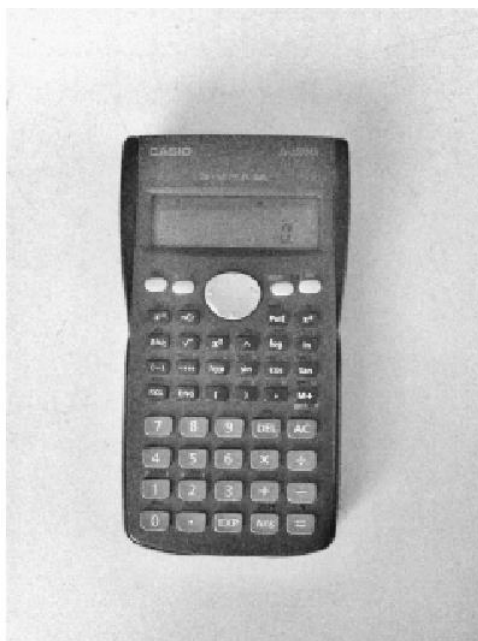
**Obrázek 20:** Podobraz "red" - s kompresí

*Zdroj [vlastní]*



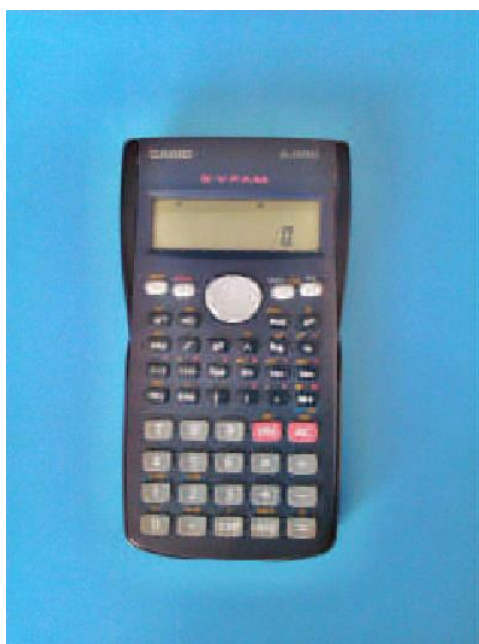
**Obrázek 21:** Podobraz "green" - s kompresí

*Zdroj [vlastní]*



**Obrázek 22:** Podobraz "blue" - s kompresí

*Zdroj [vlastní]*



**Obrázek 23:** Zkomprimovaný obraz

*Zdroj [vlastní]*

Cílem komprese je zachovat, co nejvíce významných obrazových bodů a dosáhnout, co nejmenší datové velikosti. V tomto případě bylo dosaženo 25-ti % z původní velikosti obrazu. Tedy byl splněn cíl komprese.

## ZÁVĚR

Cílem práce byla analýza dat s využitím autokorelační funkce. To bylo učiněno na třech typech dat, využívajících při zpracování autokorelační funkci. Jako první data byla zvolena časová řada, na kterou byla aplikována Box-Jenkinsova metodologie. Za druhá data byl zvolen signál, který byl vygenerován pomocí MATLABu. Byla u něho zjišťována jeho periodičita a byla na něj aplikována neuronová síť, která určila funkční zápis autokorelační funkce. Jako třetí a poslední typ dat, byly zpracovány 3 obrazy, různé velikosti a kvality.

Nejprve bylo v práci teoretické uvedení do problematiky. Byly tu popsány jednotlivé vztahy, data, metody a proměnné potřebné pro praktické zpracování.

Na časové řadě s daty udávajícími indexy spotřebitelských cen za jednotlivá čtvrtletí od roku 2000 až do roku 2013 bylo ukázáno hledání vhodného modelu řady dle Box-Jenkinsovy metodologie. Na základě provedených analýz a posouzení se došlo k závěru, že nejvhodnějším modelem této řady je AR(3). Tento model byl vygenerován v prostředí MATLAB.

Na datech uvádějících vývoj cen benzínu, nafty a LPG byla provedena analýza zjišťující výhodnost koupě těchto komodit pro hypotetickou firmu s vozovým parkem. Data byla vyhodnocena, jako nestacionární, takže byla vhodně upravena, aby se následně mohlo dojít k závěru, který byl hledán. Jako nejvýhodnější období pro koupi komodit je pro tuto firmu polovina každého měsíce.

Pro zpracování signálu byl vygenerován signál, na němž byly provedeny pozorování dle teoretického popisu ve třetí kapitole, byla zde určena na základě teoretického popisu periodičita signálu. Dalším vygenerovaným signálem byl zašuměný signál, z něhož byl následně šum odstraněn s využitím autokorelační funkce.

Byla tu ukázána Karhunen-Loeveho transformace v rámci aplikace komprese obrazu, kterým je nejprve upravené černobílé logo Univerzity Pardubice, byl na něm ukázán vliv komprese na kvalitu obrazu. Byla zde potvrzena pravdivost tvrzení, že lze s kompresí od 60-ti do 80-ti % pozorovat subjektivní zlepšení zašuměného obrazu. Jako další obraz byl použit černobílý snímek vesmíru, kde byla ukázána optimalizace výpočtu. Změna čtverců ze 40-ti obrazových bodů na 15 obrazových bodů vedla k urychlení výpočtu o 96 %, a to bez ztráty na kvalitě. Tedy cíl optimalizace výpočtu byl nad očekávání splněn. Posledním obrazem byla barevná fotografie kalkulačky, kde bylo ukázáno navíc rozložení obrazu na jednotlivé barevné vrstvy RGB. Cílem komprese bylo zachovat, co nejvíce významných obrazových bodů a

dosáhnout, co nejmenší datové velikosti. V případě fotografie kalkulačky bylo dosaženo 25-ti % z původní velikosti obrazu. Tedy byl splněn cíl komprese.

Vše bylo zpracováno za pomoci MATLAB R2010a, některé výstupy byly upraveny v MS Excel.

## POUŽITÁ LITERATURA

- [1] HEISIG, Jan. *Aplikace v regionální a sociální geografii* [online]. c2004-2013 [cit. 6. červen 2013]. Dostupné na: <<http://aplikacergsg.sci.muni.cz/teorie/trh-prace-a-ekonomika/graficke-vyjadreni-rozdeleni-cetnosti>>
- [2] NOVOVIČOVÁ, Jana. *Pravděpodobnost a matematická statistika*. ČVUT, Fakulta Dopravní 1999. [cit. 6. červen 2013]. Dostupné na: <<http://euler.fd.cvut.cz/publikace/files/skripta3.pdf>>
- [3] ARLT, Josef, ARLTOVÁ, Markéta. *Finanční časové řady*. Grada Publishing a.s. 2003, ISBN 80-247-0330-0
- [4] KUBANOVÁ, Jana. *Statistické metody pro ekonomickou a technickou praxi*. Bratislava STATIS 2003, ISBN 80-85659-31-X
- [5] TREŠL, Jiří. *Statistické metody a kapitálové trhy*. VŠE Praha 1999, ISBN 80-7079-911
- [6] TŮMA, Jiří. *Složitě systémy řízení, I. Díl: Regulace soustav s náhodnými poruchami*. 1. vyd., Ostrava 1998. 151 s. ISBN 80-7078 - 534 - 9
- [7] FRIBERT, Miroslav. *Základy zpracování obrazu*. Skriptum Univerzita Pardubice, Fakulta Chemicko-technologická 2006. ISBN 80-7194-901-9. ISBN 80-7194-901-9
- [8] KŘIVÝ, Ivan. *Analýza časových řad*. Ostrava, 2006. [cit. 6. červen 2013]. Dostupné z: <[http://informatika-osu.czechian.net/files/is/ancas/ANCAS\\_DiV.pdf](http://informatika-osu.czechian.net/files/is/ancas/ANCAS_DiV.pdf)>
- [9] RAMÍK, Jaroslav. *Statistika*. Slezská univerzita v Opavě, Obchodně podnikatelská fakulta Karviná 2007. [cit. 6. červen 2013]. Dostupné na: <<http://polodriver.uvadi.cz/files/Statistika/Statistika.pdf>>
- [10] KRIŠTOF, Aleš. *Nové metody a přístupy k analýze a prognóze ekonomických časových řad*, Praha 2006. [cit. 6. červen 2013]. Dostupné na: <<http://www.pef.czu.cz/cs/?dl=1&f=12920>>
- [11] CIPRA, Tomáš. *Analýza časových řad s aplikacemi v ekonomii*. SNTL, Praha 1986
- [12] ARLT, Josef, ARLTOVÁ, Markéta. *Ekonomické časové řady*. Grada Publishing a.s. 2007, ISBN 978-80-247-1319-9
- [13] ARLT, Josef. *Moderní metody modelování ekonomických časových řad*. Grada Publishing a.s. 1999, ISBN 80-7169-539-4

- [14] ARLT, Josef, ARLTOVÁ, Markéta, RUBLÍKOVÁ, Eva. *Analýza ekonomických časových řad s příklady*. Praha 2002. [cit. 6. červen 2013]. Dostupné na: <<http://nb.vse.cz/~arltova/vyuka/crsbir02.pdf>>
- [15] ZATLOUKAL, Radomír. *Analýza a předpověď časových řad pomocí statistických metod se zaměřením na metodu Box-Jenkins*. VUT Brno 2008. [cit. 6. červen 2013]. Dostupné z: <<https://dspace.vutbr.cz/bitstream/handle/11012/2451/diplomovaprace.pdf?sequence=1>>
- [16] HANČLOVÁ, Jana. TVRDÝ, Lubor. *Úvod do analýzy časových řad*. VŠB-TU Ostrava, 2003. [cit. 6. červen 2013]. Dostupné z: <[http://gis.vsb.cz/pan-old/Skoleni\\_Texty/TextySkoleni/AnalyzaCasRad.pdf](http://gis.vsb.cz/pan-old/Skoleni_Texty/TextySkoleni/AnalyzaCasRad.pdf)>
- [17] RUBLÍKOVÁ, E. *Analýza časových radov*. Bratislava: Ekonomická univerzita v Bratislavě, 2007. ISBN 978-80-8078-139-2
- [18] ARLTOVÁ, Markéta., ARLT, Josef. *Statistika 32: (11)*. str. 483 - 493, ČSÚ, 1995. ISSN 0322-788x.
- [19] PETRÁŠKOVÁ, Vladimíra. *Prognostické modely v oblasti modelování finančních časových řad*. ČZU Praha 2006. [cit. 6. červen 2013]. Dostupné z: <<http://www.pef.czu.cz/cs/?dl=1&f=12935>>
- [20] POPELKA, Jan, SYNEK, Václav. *Úvod do statistické analýzy dat*. Ústí nad Labem, Univerzita Jana Evangelisty Purkyně v Ústí nad Labem, Fakulta životního prostředí. 2009. 200 s. ISBN 978-80-7414-117.
- [21] BARTLETT, Maurice Stevenson. *On the theoretical specification of sampling properties of autocorrelated*. 1946
- [22] QUENOUILLE, Maurice Henry. *Approximate tests of correlation in time series*. 1949
- [23] HINDLS, Richard; HRONOVÁ, Stanislava; SEGER, Jan; FISHER, Jakub. *Statistika pro ekonomy*. Osmé vydání, Professional Publishing, 2007.
- [24] ROZINEK, Ondřej. *Modely zpracování signálů v MATLAB/Simulink-u*. Univerzita Pardubice, 2009. [cit. 6. červen 2013]. Dostupné na: <<http://dspace.upce.cz/handle/10195/35008>>
- [25] UHLÍŘ, Jan, SOVKA, Pavel. *Číslicové zpracování signálu*. ČVUT Praha 1995. ISBN 80-01-01303-0

- [26] PAVLÍK, Radomír, POLÁČEK, Vladimír. *Detekce užitečného signálu v aplikaci harmonického radaru s využitím MATLAB*. VTÚO Brno 2009. [cit. 6. červen 2013]. Dostupné na: <[http://dsp.vscht.cz/konference\\_matlab/MATLAB09/prispevky/079\\_pavlik.pdf](http://dsp.vscht.cz/konference_matlab/MATLAB09/prispevky/079_pavlik.pdf)>
- [27] VEJRAŽKA, František. *Signály a soustavy*. ČVUT Praha 1992. [cit. 6. červen 2013]. Dostupné na: <[http://home.zcu.cz/~ostember/Signaly\\_a\\_soustavy.pdf](http://home.zcu.cz/~ostember/Signaly_a_soustavy.pdf)>
- [28] KOLÁŘ, Radek. *Analýza signálu FM rozhlasu*. Univerzita Pardubice, FEI 2009. [cit. 6. červen 2013]. Dostupné na: <[http://dspace.upce.cz/bitstream/10195/33861/5/KolarR\\_AnalyzaFM\\_PB\\_1%20cast\\_2009.pdf](http://dspace.upce.cz/bitstream/10195/33861/5/KolarR_AnalyzaFM_PB_1%20cast_2009.pdf)>
- [29] ANDRLA, Petr. *Segmentace řečového signálu*. VUT Brno 2008. [cit. 6. červen 2013]. Dostupné na: <[http://www.vutbr.cz/www\\_base/zav\\_prace\\_soubor\\_verejne.php?file\\_id=8160](http://www.vutbr.cz/www_base/zav_prace_soubor_verejne.php?file_id=8160)>
- [30] FRIBERT, Miroslav. *Základy zpracování obrazu*. Univerzita Pardubice, FChT 2006. ISBN 80-71-94901-9
- [31] RABBANI, Majid, JONES, Paul W. *Digital Image Compression Techniques*. 1991, ISBN 0-8194-0648-1
- [32] ROSENFELD, Azriel., KAK, Avinash C. *Digital Picture Processing*. Academic Press, New York 1982. ISBN 0-12-597301-2
- [33] FRITSCH, Lukáš. *Simulace obrazového kodéru na bázi 3D KLT*. ČVUT Praha 2008. [cit. 6. červen 2013]. Dostupné na: <[http://phobos.vscht.cz/konference\\_matlab/MATLAB08/prispevky/032\\_fritsch.pdf](http://phobos.vscht.cz/konference_matlab/MATLAB08/prispevky/032_fritsch.pdf)>
- [34] PELÁNEK, Radek. *Modelování a simulace komplexních systémů*. Masarykova univerzita Brno 2011. ISBN 978-80-210-5318-2
- [35] CENIA, česká informační agentura životního prostředí. *Vítejte na Zemi ...*, c2013 [cit. 16. červen 2013]. Dostupné na: <[http://www.cittadella.cz/cenia/?p=index\\_spotrebitelskych\\_cen&site=spotreba](http://www.cittadella.cz/cenia/?p=index_spotrebitelskych_cen&site=spotreba)>
- [36] KURZY-online.sk, c2011-2013 [cit. 16. červen 2013]. Dostupné na: <<http://www.kurzy-online.sk/makroekonomika/benzin-nafta-lpg/>>
- [37] ČAPEK, Jan, FABIÁN, Peter. Model of the object describing with help of the Weighting function. In. *Processings of the 17<sup>th</sup> International Conference on CAD/CAM*

Robotics and Factoring of the Future, Durban, 2001 South Africa pp 1233-1238, ISBN 1-86840-432-3

- [38] ČAPEK, Jan. *Multi-Input one Model Based on the Weighting Function*. In. *Proceedings of the 10th International DAAAM Symposium, DAAAM International*, Editor B.Katalinic, Vienna, Austria, 1999. ISBN 3-901509-10-0.

## **SEZNAM PŘÍLOH**

Příloha A Index spotřebitelských cen

Příloha B Vozový park

Příloha C Generování signálu

Příloha D Zpracování obrazu

## Příloha A – Vývoj cen benzínu a nafty

Zpracovávaná data byla získána z databáze Českého statistického úřadu, ukázka získaných dat je uvedena v následující tabulce v 1. a 3. sloupci, 2. sloupec je pro zjednodušení zpracování a 4. sloupec obsahuje diferenci potřebnou pro stacionarizaci řady:

Data pro MATLAB			
Čtvrtletí	Id	CPI	1.diference
I-00	1	88,6	
II-00	2	88,9	0,3
III-00	3	89,9	1,0
IV-00	4	90,3	0,4
I-01	5	92,2	1,9
II-01	6	93,3	1,1
III-01	7	94,8	1,5
IV-01	8	94,2	-0,6
I-02	9	95,8	1,6
II-02	10	95,5	-0,3
III-02	11	95,5	0,0
IV-02	12	94,7	-0,8
I-03	13	95,4	0,7
II-03	14	95,6	0,2
III-03	15	95,4	-0,2
IV-03	16	95,5	0,1
I-04	17	97,6	2,1
II-04	18	98,1	0,5
III-04	19	98,4	0,3
IV-04	20	98,4	0,0
I-05	21	99,2	0,8
II-05	22	99,6	0,4
III-05	23	100,2	0,6
IV-05	24	100,7	0,5
I-06	25	102,0	1,3
II-06	26	102,5	0,5
III-06	27	103,2	0,7
IV-06	28	102,2	-1,0
I-07	29	103,6	1,4
II-07	30	105,0	1,4
III-07	31	105,9	0,9
IV-07	32	107,2	1,3
I-08	33	111,3	4,1
II-08	34	112,1	0,8
III-08	35	112,9	0,8

## Zdrojové kódy pro MATLAB

Kód obsahující m-file **CPI.m** obsahuje operace pro načtení vytvořeného souboru **CPI.mat**. Pro zjištění závislosti získaných časových řad je v tomto m-filu použita korelační funkce **corr()**, pro zjištění ACF a PACF pak funkce **autocorr()** a **parcorr()**.

### **CPI.m**

```
% Zpracování dat CPI
```

```
clc;
```

```
clear all;
```

```
load('CPI.mat');
```

```
n = size(CPI,1);
```

```
% Zjištění optimální hodnoty nLags
```

```
optLags = zeros(n-1,1);
```

```
for nLags = 1:1:n-1 %nLags musí být vyšší než řád vyhlazování
```

```
    optLags(nLags,1) = corr(CPI(1:end-nLags,2), CPI(nLags+1:end,2));
```

```
end
```

```
plot(optLags);
```

```
figure;
```

```
[~,nLags] = max(optLags(20:30)); % rozsah byl určen na základě grafu
```

```
optLags
```

```
nLags = nLags + 20;
```

```
[ACF,~,~] = autocorr(CPI(:,2), nLags, 3);
```

```
plot(ACF);
```

```
figure;
```

```
[PACF,~,~] = parcorr(CPI(:,2), nLags, 3);
```

```
plot(PACF);
```

Zpracování řady po diferenciaci je vidět v souboru **diference1.m**. Na konci kódu je určení reziduí a vykreslení do Q-Q grafu.

### **diference1.m**

```
% Zpracování dat CPI
```

```
clc;
```

```
clear all;
```

```
load('diference1.mat');
```

```
n = size(diference1,1);
```

```
% Zjištění optimální hodnoty nLags
```

```
optLags = zeros(n-1,1);
```

```
for nLags = 1:1:n-1 %nLags musí být vyšší než řád vyhlazování
```

```
    optLags(nLags,1) = corr(diference1(1:end-nLags,2),  
diference1(nLags+1:end,2));
```

```
end
```

```
plot(optLags);
```

```
figure;
```

```
[~,nLags] = max(optLags(25:30)); % rozsah byl určen na základě grafu
```

```
optLags
```

```
nLags = nLags + 25;
```

```
[ACF,~,~] = autocorr(diference1(:,2), nLags, 3);
```

```
plot(ACF);
```

```
figure;
```

```
[PACF,~,~] = parcorr(diference1(:,2), nLags, 3);
```

```
plot(PACF);
```

```
y = diference1(:,2);
```

```
rad = 3;
```

```
m = ar(diference1(:,2), rad);
```

```
r = resid([y(1:rad);y], m);
```

```
r = r(rad+1:end);
```

```
qqplot(r);
```

Zpracování reziduí bylo provedeno za pomoci m-filu **rezidua.m**.

### **rezidua.m**

```
% Zpracování dat CPI
```

```
clc;
```

```
clear all;
```

```
load('rezidua.mat');
```

```
n = size(rezidua,1);
```

```
% Zjištění optimální hodnoty nLags
```

```
optLags = zeros(n-1,1);
```

```
for nLags = 1:1:n-1 %nLags musí být vyšší než řád vyhlazování
```

```
    optLags(nLags,1) = corr(rezidua(1:end-nLags,1),
```

```
    rezidua(nLags+1:end,1));
```

```
end
```

```
plot(optLags);
```

```
figure;
```

```
[~,nLags] = max(optLags(1:5)); % rozsah byl určen na základě grafu optLags
```

```
[ACF,~,~] = autocorr(rezidua(:,1), nLags, 3);
```

```
plot(ACF);
```

```
figure;
```

```
[PACF,~,~] = parcorr(rezidua(:,1), nLags, 3);
```

```
plot(PACF);
```

## Příloha B – Vozový park

Zdrojový kód `pohonne_hmoty.m` načte vytvořená data a aplikuje na ně výpočet ACF, výsledky zobrazí graficky.

### `pohonne_hmoty.m`

```
clc;
clear all;

load('data');

n = size(data,1);

% Zjištění optimální hodnoty nLags
optLags = zeros(n-1,1);
for nLags = 1:1:n-1 %nLags musí být vyšší než řád vyhlazování
    optLags(nLags,1) = corr(data(1:end-nLags,2), data(nLags+1:end,2));
end
plot(optLags);
figure;
[~,nLags] = max(optLags(15:25)); % rozsah byl určen na základě grafu
optLags
nLags = nLags + 15;

[ACF,Lags,Bounds] = autocorr(data(1:n-1,2), nLags, 3);
plot(ACF);
```

## Příloha C – Generování signálu

Kódy definující zápis a definování generovaných signálů obsahují m-fily **signal1.m** a **signal2.m**.

### signal1.m

```
function [ y ] = signal1( f0, t )
    y = 0;
    f0 = 2*pi*f0;
    for i = 1:2:10
        y = y + sin(i * f0 * t) / i;
    end
end
```

### signal2.m

```
function [ y ] = signal2( f0, t )
    y = 0;
    f0 = 2*pi*f0;
    for i = 1:4:10
        y = y + sin(i * f0 * t) / i ^ 2 - sin((i + 2) * f0 * t) / (i + 2)^2;
    end
end
```

První část zdrojového kódu **signal.m** je věnována vytvoření signálu složeného ze **signal1.m** a **signal2.m**, poté následují úpravy potřebné k následnému určení periodicity signálu (průběh ACF).

Ukázka zašumění signálu je obsažena ve zdrojovém kódu **signal\_sum.m**, kde je složen jiný signál opět za pomoci **signal1.m** a **signal2.m**, narozdíl od předchozího je účelně zašuměn.

### **signal.m**

```
clc;
clear all;

t = 0:0.01:2;
t = t';
n = size(t,1);
s = zeros(n,1);
f0 = 2;

for i = 1:1:n
s(i,1) = signal1(f0, t(i,1)) + signal2(f0*2, t(i,1));
end
plot(t',s);
figure;

% Zjištění optimální hodnoty nLags
optLags = zeros(n-1,1);
for nLags = 1:1:n-1 %nLags musí být vyšší než řád vyhlazování
    optLags(nLags,1) = corr(s(1:end-nLags,1), s(nLags+1:end,1));
end
plot(optLags);
figure;
[~,nLags] = max(optLags(20:80)); % rozsah byl určen na základě grafu
optLags
nLags = nLags + 20;

[ACF,Lags,Bounds] = autocorr(s(:), nLags, 3);
plot(ACF);
```

```
figure;
```

### signal\_sum.m

```
clc;
```

```
clear all;
```

```
t = 0:0.005:2;
```

```
t = t';
```

```
n = size(t,1);
```

```
s = zeros(n,1);
```

```
f0 = 2;
```

```
for i = 1:1:n
```

```
s(i,1) = signal1(f0, t(i,1)) + signal2(f0*5, t(i,1))+rand(1);
```

```
end
```

```
plot(t',s);
```

```
figure;
```

```
% Zjištění optimální hodnoty nLags
```

```
optLags = zeros(n-1,1);
```

```
for nLags = 1:1:n-1 %nLags musí být vyšší než řád vyhlazování
```

```
    optLags(nLags,1) = corr(s(1:end-nLags,1), s(nLags+1:end,1));
```

```
end
```

```
plot(optLags);
```

```
figure;
```

```
[~,nLags] = max(optLags(75:125)); % rozsah byl určen na základě grafu
```

```
optLags
```

```
nLags = nLags + 75;
```

```
[ACF,Lags,Bounds] = autocorr(s(:), nLags, 3);
```

```
plot(ACF);
```

## Příloha D – Zpracování obrazu

Níže uvedené zdrojové kódy jsou společné pro zpracování loga, vesmíru a kalkulačky:

### autocorr.m

```
function [ acorr ] = autocorr(N,m)
% Výpočet prostorové autokorelační matice obrazu
V = N * N;
acorr = zeros(V);
for i=1:1:V
    for j=1:1:V
        k0 = mod(i-1, N) - mod(j-1, N);
        l0 = (i - j - k0) / N;
        ks = max(1, 1 + k0);
        kh = min(N, N + k0);
        ls = max(1, 1 + l0);
        lh = min(N, N + l0);
        for k = ks:1:kh
            for l = ls:1:lh
                acorr(i, j) = acorr(i, j) + (m(k, l) * m(k - k0, l - l0))/V;
            end;
        end;
    end;
end;
```

### KLtrans.m

```
function [ obraz_komprim, KL, Macf ] = KLtrans( obraz, kvalitaObrazu )

N = size(obraz, 1); % zjištění velikosti obrazu

% zjistění numerické kvality obrazu
if kvalitaObrazu > 1
    kvalitaObrazu = 1;
else if kvalitaObrazu < 0.01
    kvalitaObrazu = 0.01;
end
end
```

```

kvalitaObrazu = kvalitaObrazu*N;

vec_obraz = m2v(N,N,obraz);
mean_obraz = mean(vec_obraz);
vec_obraz_centr = vec_obraz - mean_obraz;
obraz_centr = obraz - mean_obraz;
A_obraz_centr = autocorr(N, obraz_centr);
[V,~] = eig(A_obraz_centr);
N_A = N * N;
V_rot = rotc(N_A,V);
V_sort = V_rot';
A = V_sort;

% přímá transformace
vec_KL_obraz = A * vec_obraz_centr;
vec_KL_obraz(N*kvalitaObrazu:end) = 0; % komprese
KL_obraz = v2m(N, N, vec_KL_obraz);

% inverzní transformace
obraz_inv_KL_obraz = A' * vec_KL_obraz + mean_obraz;

% sestavení výstupních hodnot
KL = KL_obraz;
Macf = v2m(N, N, A);
obraz_komprim = v2m(N,N, obraz_inv_KL_obraz);

end

```

### **m2v.m**

```

function [ v ] = m2v( M, N, m)
% Převod matice na vektor
V = M * N;
v = zeros(V, 1);
for i=1:1:N
    for j=1:1:M

```

```
        v(j + M * (i - 1),1) = m(j,i);
    end;
end;
```

#### **rotc.m**

```
function [ r ] = rotc(N, m)
% Rotace sloupců matice
r = zeros(N);
for i=1:1:N
    for j=1:1:N
        r(i, j) = m(i, N - j + 1);
    end;
end;
```

#### **v2m.m**

```
function [ r ] = rotc(N, m)
% Rotace sloupců matice
r = zeros(N);
for i=1:1:N
    for j=1:1:N
        r(i, j) = m(i, N - j + 1);
    end;
end;
```

## Zdrojový kód – zpracování loga UPce

Zápis obsažený v logo.m načítá obraz logo.png, jako matici a převádí tento obraz na stupně šedi, následně je nastavena kvalita, při níž je provedena komprese.

### **logo.m**

```
clc;
clear all;

obraz = imread('logo.png');%načtení obrázku jako matici
obraz = im2double(obraz);%konverze datového typu na double
obraz_sedy = (obraz(:,:,1)+obraz(:,:,2)+obraz(:,:,3)) / 3; % převod na
stupně šedi

imshow(obraz_sedy, 'InitialMagnification','fit');

for i = 0.6:-0.04:0
    [obraz_komprim, ~, Macf] = KLtrans(obraz_sedy, i);
    figure;
    imshow(obraz_komprim, 'InitialMagnification', 'fit');
end

%vykreslení prostorového vyjádření ACF obrazu
surf(Macf);
colormap hot;
title('Prostorové zobrazení ACF');
xlabel('X'); %X-ové souřadnice obrazu
ylabel('Y'); %Y-ové souřadnice obrazu
zlabel('ACF');
```

## Zdrojový kód – zpracování vesmíru

Obraz snímku vesmíru je zpracován dle zdrojového kódu uvedeného ve **vesmir.m**. Obraz je nejprve upraven, jako logo, následně bude rozšířen z důvodu zabránění obrazu včetně okrajových pixelů, což pak musí být po úpravě opět odříznuto.

### **vesmir.m**

```
clc;
clear all;

obraz = imread('vesmir.jpg'); % načtení obrázku jako matice
obraz = im2double(obraz); % konverze datového typu na double
obraz_sedy = (obraz(:,:,1) + obraz(:,:,2) + obraz(:,:,3)) / 3; % převod na
stupně šedi
kvalita = 0.15; % kvalita obrazu
fragmentMN = 15; % velikost čtvercových fragmentů

% rozšíření obrazu
[M N] = size(obraz_sedy);
M = M + fragmentMN;
N = N + fragmentMN;
obraz_komprim_rozsireny = zeros(M, N);
obraz_sedy_rozsireny = zeros(M, N);
obraz_sedy_rozsireny(1:size(obraz_sedy,1),1:size(obraz_sedy,2)) =
obraz_sedy;

% postupné rozřezávání a spojování obrazu
for x = 1:fragmentMN+1:M
    nextX = x + fragmentMN;
    for y = 1:fragmentMN+1:N
        nextY = y + fragmentMN;
        if nextX <= M && nextY <= N
            fragment = obraz_sedy_rozsireny(x:nextX,y:nextY); % odříznutí
            fragment = KLtrans(fragment, kvalita); % komprese fragmentu
            obraz_komprim_rozsireny(x:nextX,y:nextY) = fragment; %
připojení
        end;
    end;
end;
```

```

end;

% oříznutí rozšířeného obrazu
obraz_komprim =
obraz_komprim_rozsireny(1:size(obraz_sedy,1),1:size(obraz_sedy,2));

%zobrazení originálu a komprimovaného obrazu
imshow(obraz_sedy, 'InitialMagnification','fit');
figure;
imshow(obraz_komprim, 'InitialMagnification','fit');

```

### Zdrojový kód – zpracování kalkulačky

Jelikož se jedná o barevný obraz, liší se kód obsazený v kalkulacka.m o rozklad na jednotlivé barevné vrstvy RGB.

```

kalkulacka.m clc;
clear all;

obraz = imread('kalkulacka.png'); % načtení obrázku jako matice
obraz = im2double(obraz); % konverze datového typu na double
kvalita = 0.25; % kvalita obrazu
fragmentMN = 10; % velikost čtvercových fragmentů

% rozložení na RGB podobrazy
[M N D] = size(obraz);
obraz_komprim = zeros(M, N, D);
for d = 1:1:D
    podobraz = obraz(:,:,d);
    imshow(podobraz, 'InitialMagnification','fit');
    figure;

    % rozšíření obrazu
    M = M + fragmentMN;
    N = N + fragmentMN;
    podobraz_komprim_rozsireny = zeros(M, N);
    podobraz_rozsireny = zeros(M, N);
    podobraz_rozsireny(1:size(podobraz,1),1:size(podobraz,2)) = podobraz;

```

```

% postupné rozřezávání a spojování obrazu
for x = 1:fragmentMN+1:M
    nextX = x + fragmentMN;
    for y = 1:fragmentMN+1:N
        nextY = y + fragmentMN;
        if nextX <= M && nextY <= N
            fragment = podobraz_rozsireny(x:nextX,y:nextY); %
odřiznutí
            fragment = KLtrans(fragment, kvalita); % komprese
fragmentu
            podobraz_komprim_rozsireny(x:nextX,y:nextY) = fragment; %
připojení
                end;
            end;
        end;

% ořiznutí rozšířeného obrazu
podobraz_komprim =
podobraz_komprim_rozsireny(1:size(podobraz,1),1:size(podobraz,2));
imshow(podobraz, 'InitialMagnification','fit');
figure;
obraz_komprim(:,:,d) = podobraz_komprim;
end;

%zobrazení originálu a komprimovaného obrazu
imshow(obraz, 'InitialMagnification','fit');
figure;
imshow(obraz_komprim, 'InitialMagnification','fit');

```