

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

MODELOVANIE EKONOMICKÝCH A FINANČNÝCH
ČASOVÝCH RADOV

DIPLOMOVÁ PRÁCA

2013

Bc. Zuzana Holá

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

MODELOVANIE EKONOMICKÝCH A FINANČNÝCH
ČASOVÝCH RADOV

DIPLOMOVÁ PRÁCA

Študijný odbor: 1114 Aplikovaná matematika

Študijný program: Ekonomická a finančná matematika

Vedúca diplomovej práce: Prof. RNDr. Magda Komorníková CSc.

Bratislava, 2013

Bc. Zuzana Holá



Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Bc. Zuzana Holá
Študijný program: ekonomická a finančná matematika (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor: 9.1.9. aplikovaná matematika
Typ záverečnej práce: diplomová
Jazyk záverečnej práce: slovenský

Názov: Modelovanie ekonomických a finančných časových radov

Cieľ: Cieľom práce je modelovať ekonomické a finančné časové rady pomocou rôznych tried lineárnych a nelineárnych (najmä viacrežimových) modelov a porovnať ich popisné a predikčné vlastnosti.

Vedúci: prof. RNDr. Magda Komorníková, CSc.
Katedra: FMFI.KAMŠ - Katedra aplikovanej matematiky a štatistiky
Vedúci katedry: prof. RNDr. Daniel Ševčovič, CSc.
Dátum zadania: 25.01.2012

Dátum schválenia: 26.01.2012

prof. RNDr. Daniel Ševčovič, CSc.
garant študijného programu

študent

vedúci práce

ČESTNÉ PREHLÁSENIE

Čestne prehlasujem, že diplomovú prácu s názvom „Modelovanie ekonomických a finančných časových radov“ som vypracovala samostatne iba s použitím uvedenej literatúry a s pomocou vedúcej diplomovej práce. Svoje prehlásenie potvrdzujem podpisom.

.....

V Bratislave, apríl 2013

POĎAKOVANIE

Rada by som sa poďakovala Prof. RNDr. Magde Komorníkovej, CSc., školiteľke mojej práce, za jej rady, ochotu, pomoc a odborné vedenie pri zdolávaní problémov a hľadání odpovedí na otázky v mojej diplomovej práci. Ďalej chcem poďakovať Mgr. Jane Lenčuchovej za poskytnutie programu na tvorbu MSW modelov.

ABSTRAKT

HOLÁ, Zuzana: *Modelovanie ekonomických a finančných časových radov*. [Diplomová práca]. Univerzita Komenského v Bratislave. Fakulta matematiky, fyziky a informatiky. Katedra aplikovanej matematiky a štatistiky. Vedúca diplomovej práce: Prof. RNDr. Magda Komorníková CSc. Bratislava: UK, 2013, 90 s.

Diplomová práca sa zaoberá analýzou časových radov. Cieľom tejto práce je modelovanie ekonomických a finančných časových radov lineárnymi a nelineárnymi modelmi, porovnávanie vlastností použitých modelov a vytváranie predpovedí. Práca zahŕňa teoretický základ analýzy časových radov a prehľad potrebných modelov. Dáta sú historické hodnoty vybraných makroekonomických a finančných ukazovateľov rôznych krajín rozdelené do siedmich skupín podľa typu. Práca obsahuje hodnotenie popisných a predikčných vlastností desiatich modelov v rámci daných typov dát. V úvode práce sú vyslovené dve hypotézy, ktorých platnosť sa snažíme overiť a jednotlivé zistenia sú popísané v závere práce. Prílohou diplomovej práce sú použité príkazy na modelovanie časových radov v programovacom jazyku R aj s popisom.

Kľúčové slová: modelovanie, časové rady, analýza, lineárne modely, nelineárne modely, predpovede

ABSTRAKT

HOLÁ, Zuzana: *Modeling of economic and financial time series*. [Diploma thesis]. Comenius University in Bratislava. Faculty of Mathematics, Physics and Informatics. Department of Applied Mathematics and Statistics. Mentor of work: Prof. RNDr. Magda Komorníková CSc. Bratislava: UK, 2013, 90 p.

This diploma thesis deals with the analysis of time series. The aim of this work is modeling of economic and financial time series with linear and non-linear models, comparing the properties of used models and making predictions. The work includes theoretical basis for the analysis of time series and an overview of the necessary models. Data are selected historical values of macroeconomic and financial indicators of different countries divided into seven groups according to type. The work includes a review of descriptive and predictive properties of ten models within the given data types. In the introduction are two hypotheses, which validity we are trying to verify and particular findings are described at the end of work. The annex of the thesis is the command used for modeling the time series in programming language R.

Keywords: modeling, time series, analysis, linear models, nonlinear models, predictions

OBSAH

ÚVOD.....	10
I. TEORETICKÁ ČASŤ.....	12
1. Časové rady.....	12
1.1 História časových radov.....	12
2. Stochastické procesy a ich vlastnosti.....	14
2.1 Stacionarita.....	14
2.2 Autokorelačná funkcia (ACF).....	15
2.3 Parciálna autokorelačná funkcia (PACF).....	16
2.4 Woldova reprezentácia.....	16
2.5 Operátor posunu (Lag operator).....	17
3. Dekompozícia časového radu.....	18
3.1 Trend.....	18
3.2 Sezónna zložka.....	19
3.3 Cyklická zložka.....	19
3.4 Reziduálna zložka.....	19
4. Lineárne modely stochastických procesov.....	20
4.1 Lineárne modely stacionárnych stochastických procesov.....	20
4.1.1 Autoregresné procesy (AR).....	20
4.1.2 Procesy kľzavých priemerov (MA).....	20
4.1.3 Zmiešané ARMA modely.....	21
4.2 Lineárne modely nestacionárnych stochastických procesov.....	22
4.2.1 Procesy ARIMA.....	23
4.2.2 Procesy ARFIMA.....	24
4.3 Lineárne modely volatility.....	25
4.3.1 ARCH a GARCH modely.....	25
4.3.2 Modely IGARCH (Integrated GARCH).....	29
5. Nelineárne modely stochastických procesov.....	30
5.1 Nelineárne modely volatility - rozšírenia triedy GARCH modelov.....	30
5.2 Nelineárny aditívny autoregresný model (NAAR).....	32
5.3 Viacrežimové modely.....	33
5.3.1 Prahové autoregresné modely (TAR).....	34

5.3.2 Markov-switching autoregresný model (MSW-AR).....	36
6. Informačné kritériá a testovanie hypotéz.....	37
6.1 Akaikeho informačné kritérium AIC.....	37
6.2 Bayesovo informačné kritérium BIC.....	37
6.3 Test stacionarity dát.....	38
6.4 Test homoskedasticity.....	39
6.5 Test normality rezíduí.....	40
6.6 Testovanie nezávislosti.....	41
6.7 BDS test nezávislosti.....	42
II. APLIKAČNÁ ČASŤ.....	44
7. Výber a použitie štatistického softvéru.....	44
8. Analyzované dáta.....	46
8.1 Hrubý domáci produkt HDP.....	46
8.2 Čistý export.....	47
8.3 Inflácia.....	47
8.4 Zamestnanosť.....	48
8.5 Úroková miera.....	48
8.6 Výmenný kurz.....	49
8.7 Kompozitný indikátor systematického stresu CISS.....	49
8.8 Základné charakteristiky analyzovaných dát.....	49
9. Modelovanie časového radu hodnôt inflácie vo Fínsku.....	54
10. Hodnotenie výsledkov modelovania časových radov.....	68
10.1 Výsledky modelovania hrubého domáceho produktu HDP.....	68
10.2 Výsledky modelovania čistého exportu.....	70
10.3 Výsledky modelovania inflácie.....	71
10.4 Výsledky modelovania zamestnanosti.....	72
10.5 Výsledky modelovania úrokovej miery.....	73
10.6 Výsledky modelovania vývoju kurzov.....	75
10.7 Výsledky modelovania indikátora CISS.....	76
10.8 Hodnotenie výsledkov modelovania celkovo.....	77
ZÁVER.....	79
ZOZNAM BIBLIOGRAFICKÝCH ODKAZOV.....	81
PRÍLOHA 1.....	86

ÚVOD

Analýza časových radov má širokospektrálne využitie. Časové rady sa vyskytujú nielen v ekonómii a financiách, ale aj v stavebníctve, energetike, meteorológii a mnohých iných vedných disciplínach. Témou tejto diplomovej práce je modelovanie ekonomických a finančných časových radov. Analýzou a predpovedaním vývoja ekonomických a finančných ukazovateľov sa zaoberajú mnohí odborníci a inštitúcie. Diplomová práca poskytuje nielen prehľad najčastejšie využívaných modelov, ale aj praktické preskúmanie ich popisných a predikčných vlastností na väčšej vzorke časových radov.

Práca je formálne rozdelená do dvoch na seba nadväzujúcich častí. Prvá časť práce obsahuje všeobecný prehľad najpoužívanějších modelov časových radov v dnešnej dobe. Popísané sú nielen lineárne modely, ktoré sa bežne vyučujú na univerzitách, ale aj zložitejšie nelineárne modely. Teoretická časť začína objasnením pojmu časový rad a krátkym zhrnutím histórie vývinu časových radov. Ďalej nasleduje popis a vlastnosti stochastických procesov. Dekompozícia časového radu a príprava dát na modelovanie je vysvetlená v ďalšej kapitole. Po nej nasleduje samotné predstavenie jednotlivých lineárnych modelov časových radov a ich základné charakteristiky.

Zameranie diplomovej práce je popritom na nelineárne modely, pretože ich význam pri analýze ekonomických a finančných časových radov narastá. Mnohí odborníci používajú práve nelineárne modely. Zreteľ je na viacrežimových modeloch určených pozorovateľnými a nepozorovateľnými veličinami, ktoré sa stali v posledných rokoch veľmi populárne. V práci je popísaný postup a používanie týchto modelov, ich vysvetlenie a charakteristiky.

V poslednej kapitole teoretickej časti uvádzame rozhodovacie kritéria využívané na voľbu vhodných parametrov modelu a testy, ktoré nám slúžia na overenie hypotéz o charaktere analyzovaných dát a vhodnosti použitých modelov.

Aplikačná časť je druhou časťou tejto diplomovej práce, v ktorej už aplikujeme doterajšie poznatky a všeobecnú teóriu do praxe. Na začiatku je odôvodnenie výberu štatistického softvéru R používaného na modelovanie reálnych dát. Nasleduje popis analyzovaných dát po skupinách a základné charakteristiky jednotlivých časových radov. Ako téma naznačuje, vybrali sme si dáta ekonomického a finančného charakteru.

Obsiahnuté sú viaceré významné makroekonomické ukazovatele v rôznych krajinách Európy. Analýzu týchto historických údajov robíme pomocou modelov popísaných v teoretickej časti. Všetkých desať modelov použijeme na popis a vytvorenie predpovedí každého jedného časového radu osobitne. Aplikovanie jednotlivých modelov v sebe zahŕňa výber vhodných parametrov, často skúšaním, pomocou porovnávania rozhodovacích kritérií a testovaním rôznych hypotéz.

Rozsiahlou kapitolou aplikačnej časti je detailné vysvetlenie postupu celého modelovania na jednom príklade časového radu, zahrňujúce aj vytváranie predikcií porovnávaných so skutočnými hodnotami. Posledná kapitola obsahuje hodnotenie modelovania po jednotlivých skupinách dát pomocou vytvorených tabuliek zvolených ukazovateľov kvality popisu a vytvárania predikcií modelmi. V rámci každej skupiny hľadáme model s najlepšimi popisnými a predikčnými vlastnosťami, či už zároveň alebo osobitne, pre danú skupinu. Na konci kapitoly zhodnotíme celkové výsledky práce, získané poznatky a platnosť skúmaných hypotéz.

Vyššie je zachytená len obsahová stránka diplomovej práce. Jej cieľom je však samotné modelovanie ekonomických a finančných časových radov pomocou popísaných lineárnych a nelineárnych modelov a porovnanie ich popisných a predikčných vlastností, čo v sebe zahŕňa viac ako sa môže zdať. Nutné je oboznámiť sa a naštudovať si tematiku, vytvoriť a pochopiť potrebné príkazy v programovacom jazyku R a samotné modelovanie zahrňujúce hľadanie vhodných parametrov pre každý model u jednotlivých časových radov. A keďže počet analyzovaných časových radov je 41, práca s tým spojená je naozaj rozsiahla. Vzorka dát je väčšia preto, aby sme zistili, či sa nám podarí nájsť pre jednotlivé skupiny ekonomických a finančných časových radov jeden najlepší model. Táto hypotéza je reálne málo pravdepodobná, preto ďalšou hypotézou, ktorú sa touto diplomovou prácou snažíme overiť je, že pre jednu skupinu dát je najlepšia istá trieda modelov, či už lineárnych alebo nelineárnych, a pre inú skupinu dát je vhodné použitie inej triedy modelov. Overenie tejto hypotézy si žiada dáta z viacerých krajín a rôzne ekonomické a finančné ukazovatele. Je celkom možné, že ani jedna z týchto hypotéz sa nám nepotvrdí, čo neznamená neúspech práce. Aj takéto zistenie by bolo prínosom. Znamenalo by to, že jednotlivé dáta sa od seba odlišujú, a to nielen typovo, a je potrebné sa každému časovému radu venovať osobitne.

I. TEORETICKÁ ČASŤ

1. Časové rady

Časový rad je definovaný ako súbor kvantitatívnych pozorovaní získaných opakovaným meraním usporiadaných podľa časovej postupnosti. Vo všeobecnosti sa predpokladá, že čas je diskrétna premenná. Analýza časových radov obsahuje metódy na získavanie znalostí o základných vlastnostiach pozorovaných premenných. Modely časových radov sa používajú aj na prognózovanie, čiže predpovedanie budúcich hodnôt na základe predchádzajúcich dát. Výber vhodného modelu je kľúčovým krokom k úspešnej analýze časových radov [19].

1.1 História časových radov

Časové rady hrali dôležitú úlohu už v skorých prírodných vedách. Babylončania používali časové rady v astronómii na popis relatívnej pozície hviezd a planét a na predvídanie astronomických udalostí. Základnou myšlienkou bolo, že časový rad je možné rozložiť na konečný počet nezávislých, ale nie priamo pozorovateľných komponentov, ktoré sa vyvíjajú pravidelne. Pre tento postup je nutná existencia nezávislých faktorov, ktoré ovplyvňujú pozorované premenné. V polovici 19. storočia tento metodický prístup v astronómii začali využívať Ch. Babbage a W.S.Jevons [22] v ekonómii.

Rozklad časového radu na viacero komponentov závisiacich na určitých faktoroch však prvý raz definoval až v roku 1919 W. M. Persons [13]. Rozlišoval 4 základné zložky:

- a) dlhodobý vývoj – trend,
- b) cyklická zložka s dobou dlhšou ako jeden rok – obchodný cyklus,
- c) zložka, ktorá obsahuje vzostupy a pády do jedného roka – sezónny cyklus,
- d) zložka, ktorá obsahuje všetky pohyby, ktoré nie sú ani trendom, ani obchodným cyklom a ani sezónnou zložkou - rezíduum.

19 CHATFIELD, C. The Analysis of Time Series: An Introduction, Sixth Edition. 2003.

22 KLEIN, J. L. Statistical Visions in Time: A History of Time Series Analysis. 2005

13 FISCHER, B. Decompositions of Time Series. Comparing Direct Methods in Theory and Practice. 1995.

Klasická analýza časových radov predpokladala, že systematické komponenty, t.j. trend, obchodný cyklus a sezónny cyklus nie sú ovplyvnené náhodou a môžu byť teda zastúpené deterministickou funkciou času. Stochastický vplyv je obmedzený na rezíduá.

Od roku 1970 sa v analýze časových radov začínajú využívať výsledky a metódy teórie pravdepodobnosti a matematickej štatistiky. Nový prístup už predpokladal existenciu stochastického vplyvu na všetky zložky časových radov, nielen na rezíduá. Ruský štatistik E.E.Slutzky a britský štatistik G.U.Yule [10] na začiatku minulého storočia ukázali, že časové rady s cyklickými vlastnosťami môžu byť generované prostredníctvom váženého alebo neváženého súčtu alebo rozdielu viacerých čisto náhodných procesov. Vyvinuli modely používajúce sa dodnes: moving average procesy a autoregresné procesy. Ich praktické využitie rozšírili George E.P. Box a Gwilym M. Jenkins [5], ktorí v roku 1970 vytvorili metódy na implementáciu týchto modelov empiricky. Upustili od myšlienky rôznych komponentov a predpokladajú stochastický model pre celý proces.

10 DIEBOLD, F.X., KILIAN, L., NERLOVE, M. Time Series Analysis. 2006

5 BOX, G.E.P., JENKINS, G. Time Series Analysis: Forecasting and Control. 1970.

2. Stochastické procesy a ich vlastnosti

Množinu náhodných premenných $\{X(\omega, t), \omega \in \Omega, t \in T\}$, kde Ω je výberový priestor a $T \subset R$ je indexová množina nazývame *stochastický proces*. Diskrétne časový rad tvorí množina pozorovaní $\{x_1, x_2, \dots, x_n\}$ získaných v rovnakých časových intervaloch. Je to jedna realizácia stochastického procesu $\{X_1, X_2, \dots, X_t, \dots\}$. Počet pozorovaní sa nazýva dĺžka časového radu. V tejto práci budeme dĺžku časového radu označovať písmenom n . Stochastický proces je úplne charakterizovaný n -dimenzionálnou distribučnou funkciou. Obvykle sa zameriavame na prvé dva momenty

- **stredná hodnota:** $E[X_t]$, $t = 1, 2, 3, \dots, n$,
- **rozptyl:** $D[X_t]$, $t = 1, 2, 3, \dots, n$,

a mieru závislosti **kovarianciu:** $cov[X_i, X_j]$, $i, j = 1, 2, 3, \dots, n$, $i < j$ [8]. Tieto kovariancie sa nazývajú autokovariancie, pretože sú to kovariancie náhodných premenných toho istého procesu. Ak má stochastický proces mnohorozmerné normálne rozdelenie, jeho distribučná funkcia je podrobne popísaná jej momentmi prvého a druhého rádu.

2.1 Stacionarita

Stochastický proces $\{X_t\}_{t=1}^n$ je *striktne stacionárny*, ak pre každú podmnožinu (t_1, t_2, \dots, t_n) indexovej množiny T a každé reálne číslo $h \in R$ také, že $t_j + h \in T$, $j = 1, 2, \dots, n$ sú združené distribučné funkcie vektorov $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ a $(X_{t_1+h}, X_{t_2+h}, \dots, X_{t_n+h})$ rovnaké [15]. Keďže tento koncept je ťažko dosiahnuteľný v praxi, často sa používa kovariančná stacionarita. Stochastický proces $\{X_t\}_{t=1}^n$ je *slabo (kovariančne) stacionárny*, ak:

1. $\mu_t = E[X_t] = \mu$, $\sigma_t^2 = D[X_t] = E[(X_t - \mu_t)^2] = \sigma^2$, pre $t = 1, 2, \dots, n$
2. $\gamma(s, r) = cov(X_s, X_r) = E[(X_s - \mu_s)(X_r - \mu_r)]$ je len funkciou $(s - r)$, t.j. toho, ako sú od seba náhodné veličiny X_s a X_r v čase vzdialené a nie toho, na akom úseku časovej osi sa nachádzajú.

Ďalej budeme pod pojmom stacionárny proces rozumieť proces slabo stacionárny.

8 BROCKWELL, P.J., DAVIS, R.A. Time Series: Theory and Methods. 1986.

15 FRANSES, P. H. Time series models for business and economic forecasting. 1998.

Dôležitým príkladom takéhoto procesu je **biely šum**. Je to stacionárny proces $\{U_t\}$ s vlastnosťami:

- $E[U_t] = 0$, pre všetky $t = 1, 2, 3, \dots, n$,
- $D[U_t] = \sigma^2$, pre všetky $t = 1, 2, 3, \dots, n$,
- $\text{cov}[U_s, U_t] = 0$, pre všetky $s \neq t$ [19].

2.2 Autokorelačná funkcia (ACF)

Ak normujeme autokovariancie príslušnými rozptylmi, výsledkom je funkcia, ktorú nazývame *autokorelačná funkcia*. Pre stacionárne procesy je to funkcia nezávislá od času daná nasledovným vzorcom:

$$\rho(k) = \frac{E[(X_t - \mu)(X_{t+k} - \mu)]}{E[(X_t - \mu)^2]} = \frac{\gamma(k)}{\gamma(0)} \quad (1)$$

pre $k = 0, \pm 1, \pm 2, \dots$ a má nasledovné vlastnosti:

- $\rho(0) = 1$,
- $\rho(k) = \rho(-k)$,
- $|\rho(k)| \leq 1$ pre všetky k .

Vďaka prvým dvom vlastnostiam nám stačí používať autokorelačnú funkciu len pre $k = 0, 1, 2, \dots$. Závislosť pri časových radoch generovaných stacionárnymi stochastickými procesmi popisujeme *výberovou autokorelačnou funkciou* r_k . Tvar autokorelačnej funkcie je veľmi dôležitý, pretože identifikuje príslušný lineárny model. Významná je predovšetkým hodnota k_0 , po ktorej už môžeme autokorelačnú funkciu považovať za nulovú. Ak existuje také k_0 , že pre všetky $k > k_0$ je $\rho(k) = 0$, potom za predpokladu normality uvažovaného procesu je rozptyl výberovej autokorelačnej funkcie r_k pre $k > k_0$ daný Bartlettovou aproximáciou:

$$D(r_k) = \frac{1 + 2r_1^2 + \dots + r_{k_0}^2}{n}$$

19 CHATFIELD, C. The Analysis of Time Series: An Introduction, Sixth Edition. 2003.

Graf autokorelačnej funkcie pre $k \geq 0$ sa nazýva *korelogram*. Používa sa na testovanie rezíduí procesov, či majú vlastnosti bieleho šumu [15]. Autokorelačná funkcia bieleho šumu vyzerá nasledovne

$$\rho(k) = \begin{cases} 1, & \text{ak } k = 0 \\ 0, & \text{ak } k \neq 0 \end{cases}$$

2.3 Parciálna autokorelačná funkcia (PACF)

Koreláciu premenných X_t a X_{t-k} očistenú o vplyv premenných medzi nimi $X_{t-1}, X_{t-2}, \dots, X_{t-k+1}$ nazývame *parciálna autokorelačná funkcia* [15]. Počítame ju ako podmienenú strednú hodnotu

$$\phi_{k,k} = E[(X_t - \mu)(X_{t-k} - \mu) | X_{t-1}, X_{t-2}, \dots, X_{t-k+1}]$$

Výberová parciálna autokorelačná funkcia $r_{k,k}$ s posunutím k odhaduje parciálny regresný koeficient $\phi_{k,k}$ v autoregresii k -teho rádu

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

kde e_t je premenná nekorelovaná s náhodnými premennými X_{t-j} , $j \geq 1$.

Ak je $\phi_{k,k} = 0$ pre $k > k_0$, potom pre smerodajnú odchýlku odhadu $r_{k,k}$ platí

$$\sigma(r_{k,k}) \approx \sqrt{\frac{1}{n}}, \text{ pre } k > k_0.$$

Parciálna autokorelačná funkcia bieleho šumu je $\phi_{kk} = 0$ pre $k > k_0$.

2.4 Woldova reprezentácia

V roku 1938 H. Wold [38] dokázal, že každý stacionárny stochastický proces, ktorý neobsahuje žiadnu systematickú zložku môže byť vyjadrený ako lineárna kombinácia nekorelovaných rovnako rozdelených náhodných premenných s nulovou strednou hodnotou a konštantným rozptylom σ^2 . Takémuto procesu hovoríme *lineárny proces*.

15 FRANCES, P. H. Time series models for business and economic forecasting. 1998.

38 WOLD, H. A study in the analysis of stationary time series. 1938.

Lineárny proces môžeme teda vyjadriť v tvare

$$X = \mu + \sum_{j=0}^{\infty} \psi_j U_{t-j}, \quad (2)$$

pričom koeficienty ψ_j spĺňajú $\sum_{j=0}^{\infty} \psi_j^2 < \infty$, $\psi_0 = 1$ a $\{U_t\}_{t=-\infty}^{\infty}$ je biely šum. Toto vyjadrenie stacionárneho procesu nazývajúce sa **Woldova reprezentácia** má nasledujúce vlastnosti:

- $E[X_t] = \mu$ pre všetky $t = 1, 2, 3, \dots$,
- $D[X_t] = \sigma^2 \sum_{i=0}^{\infty} \psi_i^2$ pre všetky $t = 1, 2, 3, \dots$,
- $\gamma(k) = \text{cov}(X, X_{t+k}) = \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+k}$ pre všetky $t = 1, 2, 3, \dots$,
- $\rho(k) = \frac{\gamma(k)}{\gamma(0)} = \frac{\sum_{i=0}^{\infty} \psi_i \psi_{i+k}}{\sum_{i=0}^{\infty} \psi_i^2}$ pre $k = 0, \pm 1, \pm 2, \dots$

2.5 Operátor posunu (Lag operator)

Operátor posunu L vráti hodnotu procesu posunutú o jedno obdobie dozadu. Často sa mu hovorí aj operátor oneskorenia premennej a má nasledovné vlastnosti [37]:

- $LX_t = X_{t-1}$,
- $LX_{t-1} = L(LX_t) = L^2X_t = X_{t-2}$,
- $L^kX_t = X_{t-k}$, pre $k = \dots, -2, -1, 0, 1, 2, \dots$,
- $L^0 = 1$ je identita a $(1 - L)X_t = X_t - X_{t-1}$,
- $L^mX_{t-n} = L^m(L^nX_t) = L^{m+n}X_t = X_{t-(m+n)}$.

3. Dekompozícia časového radu

Rozklad časového radu na viacero komponentov závisiacich na určitých faktoroch prvý raz definoval v roku 1919 W. M. Persons [13]. Rozlišoval 4 základné zložky:

- a) **T - trend** – dlhodobý vývoj,
- b) **S - sezónna zložka** - pravidelné kolísanie hodnôt okolo trendu s určitou periódou,
- c) **C - cyklická zložka** - cyklické zmeny s premennou periódou,
- d) **R - reziduálna zložka** – náhodné zmeny, ktoré ostanú po odstránení predošlých troch systematických zložiek.

Dekompozícia časového radu môže byť:

- **aditívna** – hodnoty časového radu sa dajú určiť ako súčet jednotlivých zložiek:

$$X_t = T_t + S_t + C_t + R_t,$$

- **multiplikatívna** – hodnoty časového radu sa dajú určiť ako súčin hodnôt jednotlivých zložiek:

$$X_t = T_t S_t C_t R_t.$$

Časové rady spravidla obsahujú reziduálnu zložku. Okrem nej môžu, ale nemusia obsahovať aj jednu, dve alebo všetky tri systematické zložky.

3.1 Trend

Trend zachytáva dlhodobé zmeny v priemernom správaní sa časového radu, napr. dlhodobý rast alebo dlhodobý pokles. Vzniká ako dôsledok síl, ktoré systematicky pôsobia v rovnakom smere. Trend v časových radoch je možné popísať pomocou trendových funkcií a kľzavých priemerov alebo kľzavých mediánov. Modelovanie trendu pomocou trendových funkcií sa používa, ak vývoj časového radu zodpovedá určitej funkcii času, napr. lineárnej, kvadratickej, exponenciálnej, S-krivky apod. Modelovanie trendu pomocou kľzavých priemerov alebo kľzavých mediánov sa používa, ak je vývoj radu v dôsledku silného vplyvu nesystematickej zložky nerovnomerný alebo má extrémne hodnoty [13].

13 FISCHER, B. Decompositions of Time Series. 1995.

3.2 Sezónna zložka

Sezónna zložka opisuje periodické zmeny v časovom rade, ktorých perióda sa rovná určitej štandardnej jednotke času alebo jej konštantnému násobku. Zachytáva zmeny, ktoré sa pravidelne opakujú. Dĺžka periódy sa určuje pomocou korelogramu.

Rozbor sezónnej zložky môže podstatne rozšíriť naše vedomosti o zákonitostiach správania sa určitého javu a prispieť ku konštrukcii dokonalejších predpovedí uvažovaného časového radu. Ďalším dôležitým cieľom je získanie sezónne očisteného časového radu, z ktorého bola sezónna zložka odstránená alebo aspoň potlačená na maximálne možnú mieru. Sezónne očistený časový rad zbavený sezónnych a náhodných fluktuácií umožňuje efektívnejšie štúdium dlhodobých tendencií, ktorým je priebeh časového radu podriadený. Na určenie sezónnej zložky sa používajú regresné metódy založené na teórii lineárneho regresného modelu [2].

3.3 Cyklická zložka

Cyklická zložka je periodická zložka, ktorej perióda nezodpovedá kalendárnym jednotkám. Je to nepravidelná fluktuácia okolo trendu, v ktorej sa strieda fáza rastu s fázou poklesu. Odhaduje sa metódami spektrálnej analýzy.

3.4 Reziduálna zložka

Reziduálna zložka ostane v časovom rade po odstránení systematických zložiek. Je tvorená fluktuáciami v priebehu časového radu, ktoré nemajú rozpoznateľný systematický charakter. Môže sa stať, že časový rad nevykazuje pri predbežnej analýze alebo grafickom znázornení výskyt žiadnej systematickej zložky, takže sa zdá, že je tvorený len reziduálnou zložkou. Vtedy je potrebné použiť objektívny štatistický test, ktorý by túto hypotézu potvrdil alebo vyvrátil. Takéto štatistické testy nazývame *testy náhodnosti*. Používajú sa aj po dekompozícii časových radov, keď chceme overiť, či sme už z časového radu eliminovali skutočne všetky systematické zložky, t.j. či reziduá po dekompozícii sú už len náhodné fluktuácie [9].

2 ARLT, J. Moderní metody modelování ekonomických časových řad. 1999.
9 CIPRA, T. Analýza časových řad s aplikacemi v ekonomii. 1986.

4. Lineárne modely stochastických procesov

Vo všeobecnosti platí, že ktorýkoľvek časový rad x_t môžeme považovať za súčet dvoch častí. Prvá časť je to, čo vieme predpovedať použitím údajov z minulosti zhromaždených v množine Ω_{t-1} a druhá časť je to, čo nevieme predpovedať [15]. A teda X_t môžeme zapísať ako

$$X_t = E[X_t|\Omega_{t-1}] + u_t, \quad (3)$$

kde $E[\cdot | \cdot]$ označuje podmienený operátor očakávania a u_t sa nazýva nepredvídateľná časť, pre ktorú platí $E[u_t|\Omega_{t-1}] = 0$. Môžeme predpokladať, že u_t spĺňa vlastnosti bieleho šumu.

4.1 Lineárne modely stacionárnych stochastických procesov

Lineárne stacionárne procesy môžeme modelovať triedou lineárnych modelov známych ako ARMA (p, q) modely. Aparát, ktorý sa pritom používa je známy ako *Box - Jenkinsova metodológia* popísaná prvý raz v roku 1976 [5]. Táto metodológia berie za základný prvok analýzy reziduálnu zložku, tvorenú korelovanými alebo závislými náhodnými veličinami. Výhodou Box-Jenkinsovej metodológie je, že analýza sa vykonáva systematicky podľa vopred daného kľúča a obsahuje aktivity ako identifikácia modelu, odhad parametrov, overovanie modelu.

Základným predpoklad týchto modelov je:

1. Hodnota náhodnej premennej X_t v čase t závisí len na predchádzajúcich náhodných premenných (deterministická časť) a na náhodných fluktuáciách (stochastická časť).
2. Závislosť X_t na predchádzajúcich p náhodných premenných $X_{t-1}, X_{t-2}, \dots, X_{t-p}$ je lineárna.

5 BOX, G.E.P., JENKINS, G. Time Series Analysis, Forecasting and Control. 1970.
15 FRANSES, P. H. Time series models for business and economic forecasting. 1998.

4.1.1 Autoregresné procesy (AR)

Autoregresný proces rádu p AR(p) je popísaný rovnicou

$$X_t = \delta + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_p X_{t-p} + u_t, \quad (4)$$

kde $\alpha_p \neq 0$ a u_t je proces bieleho šumu, väčšinou gaussovský s nulovou strednou hodnotou a rozptylom σ_u^2 [1]. Označme nasledovnú rovnicu

$$\alpha(x) = 1 - \alpha_1 x - \alpha_2 x^2 - \dots - \alpha_p x^p$$

ako *autoregresný polynóm*. Použitím operátora posunu môžeme AR(p) proces zapísať ako

$$(1 - \alpha_1 L - \alpha_2 L^2 - \dots - \alpha_p L^p) X_t = \delta + u_t$$

alebo

$$\alpha(L) X_t = \delta + u_t. \quad (5)$$

Pri predpoklade stochastickej počiatkovej podmienky, AR(p) proces je stacionárny, ak všetky korene rovnice $1 - \alpha_1 L - \alpha_2 L^2 - \dots - \alpha_p L^p = 0$ sú v absolútnej hodnote väčšie než jedna, čiže mimo jednotkového kruhu v komplexnej rovine.

Charakteristiky AR(p) procesu:

- Stredná hodnota: $E[X_t] = \frac{\delta}{1 - \alpha_1 - \alpha_2 - \dots - \alpha_p} = \mu$
- Rozptyl: $D(X_t) = \frac{\sigma_u^2}{1 - \alpha_1 \rho(1) - \alpha_2 \rho(2) - \dots - \alpha_p \rho(p)}$
- Autokorelačná funkcia: $\rho(k) = \alpha_1 \rho(k-1) - \alpha_2 \rho(k-2) - \dots - \alpha_p \rho(k-p)$
- Parciálna autokorelačná funkcia: $\phi_{k,k} = 0$ pre $k > p$.

Použitím PACF môžeme teda identifikovať rád AR procesu.

4.1.2 Procesy kĺzavých priemerov (MA)

Procesy kĺzavých priemerov (*Moving Average procesy*) rádu q MA(q) môžu byť vo všeobecnosti zapísané rovnicou

1 ANDERSON, O. Time series analysis and forecasting: the Box-Jenkins approach. 1976.

$$X_t = \mu + u_t + \beta_1 u_{t-1} + \beta_2 u_{t-2} + \dots + \beta_q u_{t-q} \quad (6)$$

s parametrom $\beta_q \neq 0$ a procesom bieleho šumu u_t s nulovou strednou hodnotou a rozptylom σ_u^2 . Použitím operátora posunu dostávame tvar

$$X_t - \mu = (1 + \beta_1 L + \beta_2 L^2 + \dots + \beta_q L^q) u_t = \beta(L) u_t \quad (7)$$

kde $\beta(L)$ je tzv. MA polynóm.

Z rovnice (6) vidíme, že Woldova reprezentácia má koeficienty ψ_k rovné nule pre $k > q$. Preto nie je problém s konvergenciou a každý MA proces konečného rádu je stacionárny, pričom nezáleží na hodnote parametrov β_j , kde $j = 1, 2, 3, \dots, q$. Proces MA(q) je *invertovateľný*, t.j. je možné ho reprezentovať ako stacionárny AR(∞) proces, ak všetky korene MA polynómu sú v absolútnej hodnote väčšie ako 1, čiže mimo jednotkového kruhu v komplexnej rovine [2].

Charakteristiky MA(q) procesu:

- Stredná hodnota: $E[X_t] = \mu$
- Rozptyl: $D[X_t] = (1 + \beta_1^2 + \beta_2^2 + \dots + \beta_q^2) \sigma_u^2 = \gamma(0)$
- Autokorelačná funkcia: $\rho(k) = \begin{cases} \frac{\beta_k + \beta_1 \beta_{k+1} + \beta_2 \beta_{k+2} + \dots + \beta_{q-k} \beta_q}{1 + \beta_1^2 + \beta_2^2 + \dots + \beta_q^2}, & \text{ak } k \leq q \\ 0, & \text{ak } k > q \end{cases}$
- Parciálna autokorelačná funkcia: $\phi_{k,k} = 0$ pre všetky k.

Použitím ACF teda môžeme identifikovať rád MA procesu.

4.1.3 Zmiešané ARMA modely

V nasledujúcej časti popíšeme modely, ktoré vznikajú spojením autoregresných procesov AR rádu p a procesov klzavých priemerov MA rádu q . Takéto modely nazývame ARMA (p, q) modely. Navrhli ich matematici Box a Jenkins v roku 1970 [5]. Umožnia nám modelovať procesy, v ktorých sa ACF ani PACF po konečnom počte členov nevynuluje.

2 ARLT, J. Moderní metody modelování ekonomických časových řad. 1999.

5 BOX, G.E.P., JENKINS, G. Time Series Analysis, Forecasting and Control. 1970.

Vo všeobecnosti ARMA (p, q) procesy môžeme zapísať ako

$$X_t = \delta + \alpha_1 X_{t-1} + \dots + \alpha_p X_{t-p} + u_t + \beta_1 u_{t-1} + \dots + \beta_q u_{t-q}, \quad (8)$$

kde u_t je biely šum, $E[u_t] = 0$, $D[u_t] = \sigma_u^2$ a musí platiť $\alpha_p \neq 0$ a $\beta_q \neq 0$. Použitím operátora posunu dostaneme rovnicu

$$(1 - \alpha_1 L - \dots - \alpha_p L^p) X_t = (1 + \beta_1 L + \dots + \beta_q L^q) u_t \quad (9)$$

alebo

$$\alpha(L) X_t = \beta(L) u_t.$$

Aby sa faktory v oboch polynómoch nemohli vykrátiť, polynómy $\alpha(L)$ a $\beta(L)$ nemôžu mať rovnaké korene. Proces je stacionárny, ak pri stochastickej počiatkovej podmienke sú splnené podmienky stability pre AR, čiže ak polynóm $\alpha(L)$ má korene mimo jednotkového kruhu. Potom môžeme odvodiť Woldovu reprezentáciu, pre ktorú musí platiť

$$\beta(L) = \alpha(L)(1 + \psi_1 L + \psi_2 L^2 + \dots).$$

Ak všetky korene polynómu $\beta(L)$ sú mimo jednotkového kruhu, proces ARMA(p,q) je invertovateľný.

Stacionárny a invertovateľný proces ARMA(p,q) môže byť reprezentovaný procesom AR(∞) ako aj procesom MA(∞). Je teda možné vygenerovať stochastický proces s nekonečným AR a MA rádom pomocou konečného počtu parametrov [1].

4.2 Lineárne modely nestacionárnych stochastických procesov

V praxi sa veľmi často stretávame s časovými radmi generovanými nestacionárnymi stochastickými procesmi, pričom nestacionarita procesu môže byť spôsobená buď v čase meniacou sa strednou hodnotou, alebo v čase meniacim sa rozptylom. V tejto časti sa budeme zaoberať práve takýmito procesmi. Integrovaný proces rádu 1, ktorý označujeme I(1) je časový rad, ktorého prvé diferencie tvoria proces bieleho šumu.

1 ANDERSON, O. Time series analysis and forecasting: the Box-Jenkins approach. 1976.

Vo všeobecnosti proces, ktorý môžeme zapísať v tvare

$$(1 - L)^d X_t = \varepsilon_t, \quad (10)$$

kde $d = 1, 2, 3, \dots$ a ε_t je proces bieleho šumu s nulovou strednou hodnotou, nazývame **integrovaný proces rádu d** a označuje sa **I(d)**, pričom d je prirodzené číslo [27].

Uvažujme proces $X_t = X_{t-1} + \varepsilon_t$, ktorý sa nazýva proces náhodnej prechádzky (*Random Walk Process*). Je to špeciálny prípad procesu AR(1) s nulovou strednou hodnotou pre $\alpha_1 = 1$. Pomocou operátora spätného posunu L ho môžeme zapísať v tvare

$$(1 - L)X_t = \varepsilon_t.$$

Charakteristiky procesu náhodnej prechádzky:

- Stredná hodnota: $E[X_t] = \mu$ je konštantná v čase.
- Rozptyl: $D[X_t] = t\sigma_\varepsilon^2$ závisí od času t a pre $t \rightarrow \infty$ rastie neohraničene.
- Autokorelačná funkcia: $\rho(k) = \sqrt{1 - \frac{k}{t}}$ pre $t \rightarrow \infty$ a dané k konverguje k jednej.

Z uvedených charakteristík je zrejmé, že proces náhodnej prechádzky je nestacionárny proces. Typickou vlastnosťou časového radu generovaného procesom náhodnej prechádzky je cyklický priebeh. Odhady hodnôt autokorelačnej funkcie klesajú k nule veľmi pomaly. V prípade parciálnej autokorelačnej funkcie je štatisticky významný len prvý odhad, ktorý sa blíži k hodnote 1.

4.2.1. Procesy ARIMA

Ak má po transformácii integrovaného procesu pomocou diferencií d -teho rádu výsledný proces autokorelačnú (ACF) a parciálnu autokorelačnú funkciu (PACF) podobnú ACF a PACF stacionárneho a invertibilného procesu ARMA (p,q), môžeme pôvodný integrovaný proces vyjadriť v tvare

$$\alpha_p(L)(1 - L)^d X_t = \beta_q(L)\varepsilon_t. \quad (11)$$

Takto definovaný proces sa nazýva **AutoRegressive Integrated Moving-Average process ARIMA (p, d, q)**. Bol navrhnutý Boxom a Jenkinsom [5]. Vlastnosti má podobné ako

27 PRIESTLY, M.B. Nonlinear and Non-Stationary Time Series. 1988.

proces náhodnej prechádzky. Časové rady generované integrovanými procesmi sa často nazývajú *časové rady s jednotkovými koreňmi*. V praxi sa najčastejšie stretávame s procesmi I(1), resp. I(2). Ďalej budeme skúmať vlastnosti procesu I(1).

Označme $z_t = (1 - L)X_t$ a uvažujme model ARIMA (p, 1, q) v tvare

$$\alpha_p(L) z_t = \alpha_0 + \beta_q(L) \varepsilon_t.$$

Je zrejmé, že stredná hodnota z_t nie je nulová. Označme ju $E[z_t] = \mu$. Po vyjadrení parametra α_0 dostávame

$$\alpha_0 = \alpha_p(L) z_t - \beta_q(L) \varepsilon_t$$

$$\alpha_0 = \mu(1 - \sum_{j=1}^p \alpha_p).$$

Model potom môžeme vyjadriť nasledovne

$$z_t = \alpha_0 \alpha_p L^{-1} + \frac{\beta_q(L)}{\alpha_p(L)} \varepsilon_t = \mu + \psi(L) \varepsilon_t,$$

kde $\psi(L) = \frac{\beta_q(L)}{\alpha_p(L)}$. Keďže $z_t = (1 - L)X_t$, tak pre X_t platí

$$X_t = \mu + X_{t-1} + \psi(L) \varepsilon_t.$$

V prípade deterministickej počiatkovej podmienky $X(0) = x_0$ máme model

$$X_t = x_0 + \mu t + \psi(L) \sum_{j=1}^t \varepsilon_j,$$

ktorý obsahuje deterministický aj stochastický trend, ktorý je definovaný ako $S_t = \sum_{j=1}^t \varepsilon_j$.

Stredná hodnota procesu X_t generujúceho časový rad x_t je potom $E[X_t] = x_0 + \mu t$.

4.2.2. Procesy ARFIMA

Zovšeobecnením procesov ARIMA (p,d,q) sú procesy **AutoRegressive Fractionally Integrated Moving-Average process ARFIMA (p,d,q)** [30]. Integrovaný proces je vyjadrený taktiež v tvare

5 BOX, G.E.P., JENKINS, G. Time Series Analysis, Forecasting and Control. 1970.

30 SWANSON, N. R., BHARDWAJ, G. An Empirical Investigation of the Usefulness of ARFIMA Models for Predicting Macroeconomic and Financial Time Series. 2006.

$$\alpha_p(L)(1-L)^d X_t = \beta_q(L)\varepsilon_t, \quad (12)$$

avšak v prípade ARFIMA (p,d,q) procesov je parameter d definovaný na intervale $(-1; 1)$. Proces je kovariančne stacionárny len pre hodnoty $d \in (-0,5; 0,5)$. Navyše, ak všetky korene polynómov $\alpha(L), \beta(L)$ sú mimo jednotkového kruhu, proces ARFIMA (p, d, q) je stacionárny a invertovateľný. Stredná hodnota procesu X_t generujúceho rad x_t je nulová.

4.3 Lineárne modely volatility

Modely volatility, ktoré navrhol Engle [12] sú charakteristické tým, že podmienený rozptyl je lineárnou funkciou štvorcov rezíduí, preto sa tieto modely a modely z nich odvodené označujú ako *lineárne modely volatility*.

4.3.1 ARCH a GARCH modely

Praktické použitie modelov ARMA môže byť postačujúcim nástrojom na analýzu lineárnych stacionárnych časových radov, v ktorých pozorovanie v čase t závisí na minulých pozorovaniach lineárne a nepodmienený aj podmienený rozptyl je konštantný. V mnohých časových radoch však môže rozptyl náhodnej premennej v čase t závisieť na minulých pozorovaniach.

Časové rady, ktorých dlhodobý nepodmienený rozptyl je konštantný, ale sú obdobia, kedy je relatívne vysoký, sa nazývajú **podmienené heteroskedastické**. Jedným z modelov, v ktorom podmienený rozptyl v čase t závisí na minulých pozorovaniach, je model časového radu s *autoregresnou podmienenou heteroskedasticitou*, model ARCH a jeho zovšeobecnenie model GARCH.

Stochastický proces $\{X_t\}$, pre ktorý platí

$$\text{cov}(X_t, X_s) = \begin{cases} \sigma^2, & \text{pre } t = s \\ 0, & \text{pre } t \neq s' \end{cases}$$

kde $t, s \in (-\infty, +\infty)$, sa nazýva **náhodný proces s homoskedasticitou**.

12 ENGLE, R. F. Autoregressive Conditional Heteroskedasticity With Estimates of the Variance of U.K. Inflation. 1982.

Stochastický proces $\{X_t\}$, pre ktorý platí

$$\text{cov}(X_t, X_s) = \begin{cases} \sigma^2(t), & \text{pre } t = s \\ 0, & \text{pre } t \neq s \end{cases},$$

teda rozptyl $D[X_t]$ je funkciou času a $t, s \in (-\infty, +\infty)$, sa nazýva **náhodný proces s heteroskedasticitou** [38].

Nech časový rad x_t je generovaný procesom X_t v tvare

$$X_t = E[X_t | \Omega_{t-1}] + u_t, \quad (13)$$

kde Ω_{t-1} je informačná množina obsahujúca všetky relevantné informácie až do času $t - 1$. Predikovateľná časť $E[X_t | \Omega_{t-1}]$ je modelovaná lineárnymi ARMA (p, q) modelmi a u_t je nepredikovateľná časť s vlastnosťami

- $E[u_t | \Omega_{t-1}] = 0$,
- $E[u_t^2] = \sigma^2$,
- $E[u_t^2 | \Omega_{t-1}] = h_t$ pre nezápornú funkciu $h_t = h_t(\Omega_{t-1})$.

Vhodný tvar na vyjadrenie u_t je napríklad

$$u_t = v_t \sqrt{h_t}, \quad (14)$$

kde v_t sú nezávislé, rovnako rozdelené náhodné premenné (proces typu i.i.d.) s nulovou strednou hodnotou $E[v_t] = 0$ a jednotkovým rozptylom $D[v_t] = 1$.

Nech u_t je teda časový rad s rozptylom $D[u_t] = \sigma^2$ vyjadrený v tvare (14), pričom pre funkciu h_t platí

$$h_t = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \dots + \alpha_m u_{t-m}^2$$

kde $\alpha_0 > 0$, $\alpha_i \geq 0$, $i = 1, 2, \dots, m$ a $\{v_t\}$ je biely šum nezávislý na h_{t-k} , pre ktorý platí $E[v_t] = 0$ a $D[v_t] = 1$. Potom u_t je časový rad s **autoregresnou podmienenou heteroskedasticitou rádu m** (*AutoRegressive Conditionally Heteroskedastic*) a označujeme ho ARCH(m) [12].

38 WHITE, H. A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. 1980

12 ENGLE, R. F. Autoregressive Conditional Heteroskedasticity With Estimates of the Variance of U.K. Inflation. 1982.

Uvažujme teraz o vlastnostiach časového radu u_t daného vzťahom (14). Ak $\alpha_j \geq 0$, $j = 1, \dots, m$ a platí $\alpha_0 + \alpha_1 + \alpha_2 + \dots + \alpha_m < 1$, časový rad $\{u_t^2\}$ je kovariančne stacionárny. Pretože v_t je biely šum a nezávisí od $u_{t-1}, u_{t-2}, \dots, u_{t-m}$, platia vzťahy:

- $E[u_t] = 0$
- $E[u_t, u_{t-k}] = 0$ pre $k > 0$.

To znamená, že prvky časového radu u_t majú nulovú strednú hodnotu a sú nekorelované. Pre rozptyl platí vzťah $D[u_t] = \frac{\alpha_0}{1 - \alpha_1 - \alpha_2 - \dots - \alpha_m}$. Podmienená stredná hodnota je rovná nule:

$$E[u_t | u_{t-1}, u_{t-2}, \dots, u_{t-m}] = E[v_t] \sqrt{\alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \dots + \alpha_m u_{t-m}^2} = 0.$$

Avšak podmienený rozptyl už rovný nule nie je:

$$D[u_t | u_{t-1}, \dots, u_{t-m}] = E[u_t^2 | u_{t-1}, \dots, u_{t-m}] = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \dots + \alpha_m u_{t-m}^2.$$

Alternatívna forma ARCH(m) modelu pre časový rad u_t , zapísaná pomocou AR(m) reprezentácie s nenulovou strednou hodnotou pre u_t^2 je:

$$u_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \dots + \alpha_m u_{t-m}^2 + w_t,$$

kde $w_t = u_t^2 - h_t$.

Nech $\{u_t\}$ je časový rad s rozptylom $D[u_t] = \sigma^2$, ktorý možno vyjadriť v tvare

$$u_t = v_t \sqrt{h_t},$$

$$h_t = \alpha_0 + \sum_{i=1}^p \delta_i h_{t-i} + \sum_{i=1}^q \alpha_i u_{t-i}^2 \quad (15)$$

kde v_t je biely šum nezávislý na h_{t-k} , $k > 0$ so strednou hodnotou $E[v_t] = 0$ a rozptylom $D[v_t] = 1$. Takto definovaný rad $\{u_t\}$ sa nazýva **časový rad so zovšeobecnenou autoregresnou podmienenou heteroskedasticitou** rádu p, q (*Generalized ARCH*) a označujeme ho **GARCH (p, q)**. Tento model bol navrhnutý Bollerslevom v roku 1986 [4].

Pretože $\{v_t\}$ je biely šum, ktorý je nezávislý na minulých realizáciách u_{t-k} , $k > 0$, podobne ako pri ARCH modeloch možno ukázať, že podmienená aj nepodmienená stredná hodnota u_t je rovná 0 a podmienený rozptyl je $D[u_t | u_{t-1}, \dots, u_{t-m}] = h_t$.

4 BOLLERSLEV, T. Generalized Autoregressive Conditional Heteroskedasticity. 1986.

Označme $r = \max\{p, q\}$. Aby mala reprezentácia GARCH (p, q) modelu zmysel, musia byť splnené podmienky [23]:

1. $\alpha_0 > 0$,
2. $\alpha_j \geq 0$ pre $j = 1, 2, \dots, r$,
3. $\delta_j \geq 0$, pre $j = 1, 2, \dots, r$.
4. $(\delta_1 + \alpha_1) + (\delta_2 + \alpha_2) + \dots + (\delta_r + \alpha_r) < 1$.

Ak sú tieto podmienky splnené, časový rad u_t^2 je kovariančne stacionárny a pre nepodmiernený rozptyl u_t dostávame vzťah

$$D[u_t] = E[u_t^2] = \frac{\alpha_0}{1 - \sum_{i=1}^p \delta_i - \sum_{i=1}^q \alpha_i}.$$

Teda GARCH proces je stacionárny, keď stredná hodnota, nepodmiernený rozptyl a autokovariancia sú konečné a konštantné v čase. To platí pri splnení podmienky

$$\sum_{i=1}^p \delta_i + \sum_{i=1}^q \alpha_i < 1.$$

4.3.2 Integrované GARCH modely IGARCH

Pri používaní GARCH modelov pre modelovanie finančných časových radov sa často stáva, že súčet odhadov parametrov $\sum_{i=1}^p \delta_i + \sum_{i=1}^q \alpha_i \approx 1$. Preto Engle a Bollerslev v roku 1986 navrhli novú triedu modelov, ktorú nazvali **Integrované GARCH modely** (*Integrated GARCH models*) **IGARCH modely** [4]. Proces IGARCH môžeme vyjadriť v tvare

$$(1 - \alpha_p(L) - \beta_q(L))u_t^2 = \alpha_0 + (1 - \beta_q(L))w_t, \quad (16)$$

kde $w_t = u_t^2 - h_t$ a $(1 - \alpha_p(L) - \beta_q(L)) \equiv \phi(L)(1 - L)$, pričom korene rovníc $\phi(L) = 0$ a $(1 - \beta_q(L)) = 0$ ležia mimo jednotkového kruhu v komplexnej rovine.

Na rozdiel od procesu náhodnej prechádzky, kde autokorelačná funkcia konvergovala k hodnote jeden, v modeli IGARCH exponenciálne klesá. Procesy IGARCH nie sú kovariančne stacionárne, ale sú striktne stacionárne.

23 KOZAK, J., ARLT, J., HINDLS, R. Úvod do analýzy ekonomických časových rad. 1994

4 BOLLERSLEV, T. Generalized Autoregressive Conditional Heteroskedasticity. 1986.

5. Nelineárne modely stochastických procesov

Ekonomovia v mnohých ekonomických a finančných časových radoch s dostatočnou časovou dĺžkou zaznamenali zmeny, ktoré náhle menia charakter doterajšieho fungovania. Pre tieto zmeny je charakteristické, že sa dejú príležitostne. Príkladmi takýchto udalostí sú svetová finančná kríza, kedy sa menové kurzy či ekonomiky viacerých krajín dramaticky prepádajú z obdobia expanzie do období recesie, vývoj cien komodít reaguje na zmeny vyplývajúce z vonkajších geopolitických udalostí a pod. Týmito vplyvmi dochádza k zmene charakteristik pozorovaného časového radu alebo k zmene koeficientov modelu. Pri takýchto časových radoch si už nevystačíme len s lineárnymi modelmi. Preto sa do pozornosti stále viac dostávajú nelineárne modely, napr. nelineárne GARCH modely, bilineárne modely, viacrežimové modely a pod. Práve veľké množstvo tried nelineárnych modelov spôsobuje, že pri nelineárnych časových radoch je veľkým problémom výber toho najlepšieho modelu. V tejto práci sa zameriame na modely, ktoré sú dobre interpretovateľné.

Pri tvorbe nelineárnych modelov sa vo všeobecnosti prijíma zásada “od špeciálneho ku všeobecnému”. Vhodné je začínať jednoduchšími lineárnymi modelmi a potom, keď nie sú splnené podmienky pre ich použitie, prejsť k modelom všeobecnejším a tiež komplexnejším, akými sú modely nelineárne. Odporúčajú sa nasledovné kroky:

1. Pre daný časový rad určiť vhodný lineárny model typu AR(p).
2. Testovať nulovú hypotézu linearity modelu oproti alternatívnej hypotéze nelinearity modelu.
3. Odhadnúť parametre zvoleného nelineárneho modelu.
4. Overiť vhodnosť modelu diagnostickými testami.
5. Ak je potrebné, modifikovať model.
6. Použiť model na popisné účely a predikciu.

5.1 Nelineárne modely volatility - rozšírenia triedy GARCH modelov

Lineárne modely volatility nie sú schopné reflektovať rôzne typy asymetrie, napríklad tzv. „leverage effect“, čo označuje jav, kedy sa kladné a záporné šoky do

podmieneneho rozptylu nepremietajú symetricky, pretože podmienený rozptyl v nich závisí len na druhých mocninách šokov, čiže kladné aj záporné šoky majú rovnaký efekt. Preto bolo navrhnutých mnoho nelineárnych modelov volatility, ktoré sú schopné túto asymetriu, ale aj iné, dobre popísať.

Jedným z často používaných modelov je **Glosten – Jagannathan – Runkle GARCH model (GJR-GARCH model)**, navrhnutý tromi matematikmi L.R.J. Glostenom, R. Jagannathanom a D.E. Runklem v roku 1993, ktorý rieši problém asymetrie pomocou indikačnej funkcie [16]. Model predpokladá špecifickú parametrickú formu podmienenej heteroskedasticity. Časový rad X_t je generovaný procesom v tvare

$$X_t = \mu_t + \varepsilon_t$$

kde $\varepsilon_t = \sigma_t z_t$, pričom z_t má nulovú strednú hodnotu a rozptyl rovný 1 a pre σ_t platí vzťah

$$\sigma_t^2 = \left(\omega + \sum_{j=1}^m \zeta_j v_{jt} \right) + \sum_{j=1}^q (\alpha_j + \gamma_j I_{t-j}) \varepsilon_{t-j}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2,$$

kde

$$I_{t-j} = \begin{cases} 0, & \text{ak } \varepsilon_{t-j} \geq 0 \\ 1, & \text{ak } \varepsilon_{t-j} < 0 \end{cases}$$

a $\alpha_j, \beta_j, \gamma_j, \zeta_j, \omega$ sú konštanty. Parameter γ_j reprezentuje „leverage“ člen.

Ďalším rozšírením triedy GARCH modelov je **Asymmetric power ARCH model (APARCH model)** navrhnutý Dingom v roku 1993 [11]. Tento model sa často používa na modelovanie volatility výnosov z akcií, pretože dobre vyjadruje vyššie spomínané asymetrie. V tomto modeli pre nepodmienенý rozptyl σ_t platí

$$\sigma_t^\delta = \omega + \sum_{i=1}^q \alpha_i (|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^\delta + \sum_{i=1}^p \beta_i \sigma_{t-i}^\delta, \quad (17)$$

kde $\delta \in R^+$. Z APARCH modelu môžu byť odvodené rôzne submodely. Keď $\delta = 2$, $\beta_i = 0$, $\gamma_i = 0$ pre $i = 1, \dots, \max(p, q)$, dostávame tradičný ARCH model. Ďalej keď $\delta = 2$ a $\gamma_i = 0$, dostávame štandardný GARCH model. Nakoniec, keď $\delta = 1$, dostávame vyššie

16 GLOSTEN, L. R. J., JAGANNATHAN, R., RUNKLE, D. E., On The Relation between the Expected Value and The Volatility of Nominal Excess Return on stocks. 1993.

11 DING, Z., GRANGER, C.V.J., ENGLE, R.F. A long memory property of stock market returns and a new model. 1993.

spomínaný GJR-GARCH model. Ak vo vzorci (17) dosadíme $\delta = 1$ a $\gamma_i = 0$ pre $i = 1, 2, \dots, q$, dostaneme **Absolute Value GARCH model (AVGARCH model)** navrhnutý Taylorom v roku 1986 a Schwertom v roku 1990 [29].

Ďalším odvodením od APARCH modelu je aj **Threshold GARCH model (TGARCH model)** navrhnutý Zakoianom v roku 1994 [42], v ktorom pre nepodmienený rozptyl σ_t platí

$$\sigma_t^2 = \left(\omega + \sum_{j=1}^m \zeta_j v_{jt} \right) + \sum_{j=1}^q (\alpha_j \sigma_{t-1} (|\varepsilon_{t-j}| - \eta_{1,j} \varepsilon_{t-j})) + \sum_{j=1}^p \beta_j \sigma_{t-j}^2,$$

kde $|\eta_{1,j}| \leq 1$ pre $j = 1, 2, \dots, q$.

Existujú ešte mnohé iné modifikácie nelineárnych GARCH modelov, ktoré však vo svojej práci už nebudeme používať.

5.2 Nelineárny aditívny autoregresný model (NAAR)

Časový rad x_t môžeme modelovať nelineárnym aditívnym autoregresným modelom (NAAR), ktorý má tvar

$$X_t = f_0 + \sum_{i=1}^p f_i(X_{t-i}) + u_t, \quad (18)$$

kde f_i , je jednorozmerná spojitá funkcia pre $i = 0, 1, \dots, p$. Pretože každá funkcia f_i má jeden argument, na jej neparametrický odhad je možné použiť jednorozmernú techniku vyhladenia, ktorú popisujú Tsay a Chen (1993) [36]. Tento model je veľmi flexibilný, pretože zahŕňa lineárne modely AR a oveľa zaujímavejšie nelineárne modely v špeciálnych prípadoch. Model zovšeobecňuje lineárne regresné modely a umožňuje interpretáciu marginálnych zmien, ako napríklad vplyv niektorej premennej na strednú hodnotu. Navyše veľkou výhodou NAAR modelu je, že relatívne presný odhad dosahuje aj pre relatívne krátke časové rady.

29 SCHWERT, G.W. Stock volatility and the crash of '87. 1990.

42 ZAKOIAN, J.M. Threshold heteroskedastic models. 1994.

36 TSAY, R., CHEN, R. Nonlinear additive ARX models, Journal of the American Statistical Association. 1993.

5.3 Viacrežimové modely

Pre ekonomické a finančné časové rady sú typické nelineárne vlastnosti. Dôležitým príkladom je existencia viacerých režimov, v ktorých výnosy a volatilita majú dynamické správanie. Čiže prirodzeným prístupom k modelovaniu finančných časových radov je definovať viacero režimov, kde dynamické správanie premenných závisí na režime, v ktorom sa vyskytuje v danom okamihu. Takýmto dynamickým správaním sa rozumie, že niektoré vlastnosti časových radov, ako napr. stredná hodnota, rozptyl, autokorelácia sa v jednotlivých režimoch líšia. Príkladom môže byť, že každý kvartál bude časový rad v inom režime. Takýto jav sme popisovali pojmom sezónnosť dát. V prípade sezónnosti je zmena režimu deterministický proces, t.j. predpokladá sa, že režim, v ktorom sa v určitom čase budú dáta nachádzať je s určitosťou známy vopred. Viacrežimové modely však interpretujú režim ako stochastický proces. Čiže nevieme s určitosťou povedať, aký režim bude nasledovať, kedy nastane zmena a pri niektorých typoch viacrežimových modelov dokonca nevieme ani to, aký režim aktuálne je. Takýto stochastický proces je relevantný pre ekonomické a finančné dáta. LeBaron v roku 1992 [24] ukázal, že autokorelácia výnosov z akcií súvisí s úrovňou volatility týchto výnosov. Teda autokorelácia je väčšia počas periódy s nízkou volatilitou a nižšia počas periódy s vysokou volatilitou. Períodu nízkej a periódu vysokej volatility môžeme chápať ako dva rôzne režimy.

Ďalej popíšeme nelineárne modely, v ktorých budeme na popis dynamického správania sa časového radu v jednotlivých režimoch používať AR modely. Inými slovami časový rad je modelovaný AR modelmi, avšak autoregresné parametre závisia od režimu. Dostupné viacrežimové modely sa líšia v spôsobe, akým sa režim prepína v čase.

Prvá skupina modelov je taká, kde režim možno charakterizovať alebo stanoviť pozorovateľnou premennou. V týchto modeloch sa predpokladá, že režimy, ktoré sa objavili v minulosti a súčasnosti sú známe, resp. je ich možné identifikovať štatistickými technikami. Modely v druhej triede predpokladajú, že režim prakticky nemožno pozorovať, ale je určený podkladovým stochastickým procesom. V týchto m -režimových modeloch je zmena režimu určená stavovou premennou s_t , nadobúdajúcou len hodnoty $1, 2, \dots, m$, ktorú reprezentuje m -stavový Markovov proces prvého rádu s prechodovou maticou P .

24 LEBARON, B. Forecast improvements using a volatility index. 1992.

Jednotlivé režimy teda nie je možné identifikovať presne, ale len s určitou pravdepodobnosťou.

5.3.1 Prahové autoregresné modely (TAR)

Základným modelom prvej skupiny je **Prahový autoregresný model** (*Threshold Autoregressive model*) **TAR** navrhnutý Tongom v roku 1978 [33]. V modeloch TAR sa vhodný režim v čase t určuje podľa vzťahu pozorovateľnej indikačnej premennej q_t k prahovej (*threshold*) hodnote, označovanej ako c . Indikačná premenná q_t môže byť hodnota iného časového radu, modelujúceho ukazovateľ ovplyvňujúci hodnoty sledovaného časového radu, resp. hodnota sledovaného časového radu posunutá v čase o d časových intervalov, teda $q_t = X_t - d$ pre určité prirodzené číslo d . Takýto model sa nazýva **Self-Exciting Threshold Autoregressive model SETAR**.

Napríklad 2-režimový SETAR model s režimami popísanými $AR(p_1)$ a $AR(p_2)$ má tvar

$$X_t = \begin{cases} \phi_{1,0} + \phi_{1,1}X_{t-1} + \dots + \phi_{1,p_1}X_{t-p_1} + \varepsilon_t, & \text{ak } X_{t-d} \leq c \\ \phi_{2,0} + \phi_{2,1}X_{t-1} + \dots + \phi_{2,p_2}X_{t-p_2} + \varepsilon_t, & \text{ak } X_{t-d} > c \end{cases} \quad (19)$$

kde $\{\varepsilon_t\}$ je proces bieleho šumu s nulovou strednou hodnotou a rozptylom σ_ε^2 .

Iný spôsob ako zapísať takýto SETAR model je pomocou indikačnej funkcie:

$$\begin{aligned} X_t &= Y_{t,1}'\phi_1 I(X_{t-d} \leq c) + Y_{t,2}'\phi_2 I(X_{t-d} > c) + \varepsilon_t \\ &= Y_{t,1}'\phi_1 [1 - I(X_{t-d} > c)] + Y_{t,2}'\phi_2 I(X_{t-d} > c) + \varepsilon_t, \end{aligned}$$

kde $Y_{t,i} = (1, X_{t-1}, X_{t-2}, \dots, X_{t-p_i})'$ a $\phi_i = (\phi_{i,0}, \phi_{i,1}, \dots, \phi_{i,p_i})'$, pričom $i = 1, 2$.

Ak funkciu $I[q_t > c]$ nahradíme spojitou funkciou $G(q_t; \gamma; c)$ s *parametrom vyhladzovania* γ , ktorý sa hladko mení z 0 na 1, dostaneme model s názvom **Smooth Transition Autoregressive model STAR** [15]. Napríklad 2-režimový STAR model s režimami popísanými $AR(p_1)$ a $AR(p_2)$ má tvar

$$X_t = \phi_1' Y_t + (\phi_2 - \phi_1)' Y_t G(q_t; \gamma; c) + \varepsilon_t, \quad (20)$$

33 TONG, H. On a threshold model. Patternrecognition and signalprocessing. 1978.
15 FRANSES, P.H., DIJK, D. Nonlinear Time Series Models in Empirical Finance. 2003.

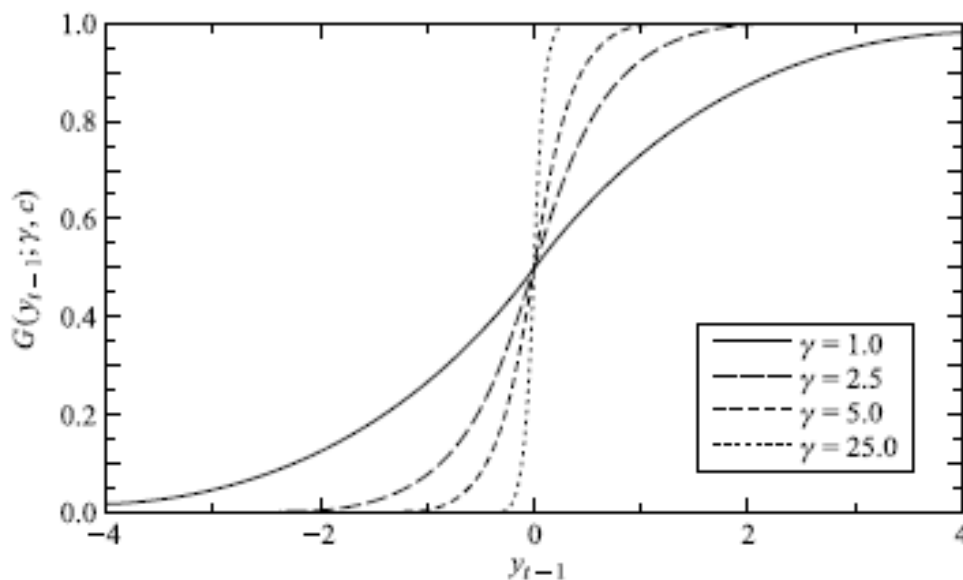
Kde $p = \max(p1, p2)$, $Y_t = (1, X_{t-1}, X_{t-2}, \dots, X_{t-p})'$ a $\phi_i = (\phi_{1,0}, \phi_{1,1}, \dots, \phi_{1,p})'$, pričom $i = 1, 2$ a $\{\varepsilon_t\}$ je proces bieleho šumu s nulovou strednou hodnotou a rozptylom σ_ε^2 .

Funkcia $G(q; \gamma; c)$ sa nazýva *prechodová funkcia*. Parameter vyhladzovania γ určuje *hladkosť* zmeny hodnôt prechodovej funkcie a teda aj rýchlosť prechodu z jedného režimu do druhého.

Často používanou prechodovou funkciou je logistická funkcia:

$$G(q_t; \gamma; c) = \frac{1}{1 + \exp(-\gamma[q_t - c])} \quad (21)$$

Pri použití prechodovej funkcie v tvare (21) sa model nazýva **Logistický STAR model** (*Logistic STAR model*) **LSTAR** [15]. Parameter c môže byť interpretovaný ako rozhranie medzi režimami zodpovedajúcimi hodnotám funkcie $G(q_t; \gamma; c) = 1$ a $G(q_t; \gamma; c) = 0$. Navyše platí, že $G(c; \gamma; c) = 0,5$. Na obrázku č. 2 sú zobrazené príklady logistickej funkcie $G(y_{t-1}; \gamma; 0)$ pre rôzne hodnoty parametra γ . Je vidieť, že ak je parameter γ veľmi veľký, zmena funkcie $G(q_t; \gamma; c)$ z 0 na 1 je takmer okamžitá. Keď sa parameter γ blíži k ∞ , dostávame SETAR model. V prípade, že $\gamma = 0$, dostávame lineárny model.



Obrázok č. 2 : Príklady logistickej funkcie $G(y_{t-1}; \gamma; 0)$ pre rôzne hodnoty parametra γ [15].

5.3.2 Markov-Switching Autoregresný model (MSW-AR)

Druhá trieda viacrežimových modelov predpokladá, že režim, ktorý sa vyskytuje v čase t nie je možné pozorovať, ale stanovuje sa stochastickým procesom, ktorý označujeme ako s_t . Aby bol proces plne popísaný, treba špecifikovať vlastnosti procesu s_t . Najznámejším modelom v tejto skupine je **Markov-Switching Autoregresný model MSW-AR**, ktorý navrhol Hamilton v roku 1989 [17]. Proces s_t je homogénny diskretný Markovovský stochastický proces (Markovov reťazec) prvého stupňa. To znamená, že režim v čase t závisí len na predchádzajúcom režime v čase $t - 1$. Model dopĺňajú pravdepodobnosti prechodu z jedného režimu do druhého. Pravdepodobnosť p_{ij} je rovná pravdepodobnosti, že Markovov reťazec prejde zo stavu i v čase $t - 1$ do stavu j v čase t . Inak povedané, je to pravdepodobnosť, že režim i v čase $t - 1$ je nasledovaný režimom j v čase t . Môžeme ju zapísať ako

$$\Pr(s_t = j | s_{t-1} = i, s_{t-2} = k, \dots) = \Pr(s_t = j | s_{t-1} = i) = p_{i,j}.$$

Pre m -režimový MSW-AR model platí, že $p_{i,j} \geq 0$ pre $i, j = 1, 2, \dots, m$, $\sum_{j=1}^m p_{i,j} = 1$ pre $i = 1, 2, \dots, m$.

Napríklad 2-režimový MSW-AR proces s $AR(p_1)$ a $AR(p_2)$ v jednotlivých režimoch je v tvare

$$X_t = \begin{cases} \phi_{1,0} + \phi_{1,1}X_{t-1} + \dots + \phi_{1,p_1}X_{t-p_1} + \varepsilon_t, & \text{ak } s_t = 1 \\ \phi_{2,0} + \phi_{2,1}X_{t-1} + \dots + \phi_{2,p_2}X_{t-p_2} + \varepsilon_t, & \text{ak } s_t = 2 \end{cases}, \quad (22)$$

kde $\{\varepsilon_t\}$ je proces typu i.i.d s nulovou strednou hodnotou. Tento model môžeme zapísať aj v kratšej forme

$$X_t = \phi_{s_t,0} + \phi_{s_t,1}X_{t-1} + \phi_{s_t,2}X_{t-2} + \dots + \phi_{s_t,p}X_{t-p} + \varepsilon_t,$$

kde $p = \max(p_1, p_2)$, $s_t = 1, 2$.

Výhodou MSW-AR modelov oproti prahovým TAR modelom je, že nie je nutná *a priori* voľba indikačnej premennej.

17 HAMILTON, J. D., A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. 1989

6. Informačné kritériá a testovanie hypotéz

Často sa stáva, že len podľa dát, ich grafického zobrazenia, ACF alebo PACF nevieme spoľahlivo určiť rád modelu. Na identifikáciu vhodného rádu modelu sa preto používajú informačné kritériá. Sú založené na hľadaní kompromisu medzi príliš vysokým rádom modelu, ktorý znižuje rozptyl reziduálnej zložky, a príliš nízkym rádom modelu, čoho dôsledkom je vysoká hodnota rozptylu reziduálnej zložky. Lepší model je ten, pre ktorý je menšia hodnota informačného kritéria.

Existuje veľa typov informačných kritérií. Uvedieme si dve základné, ktoré budeme v praxi používať.

6.1 Akaikeho informačné kritérium AIC

Jedno z najpoužívanejších kritérií je Akaikeho informačné kritérium AIC, ktoré navrhol japonský štatistik Hirotugu Akaike v roku 1973 [41]. Myšlienkou AIC je výber parametrov modelu tak, aby sa minimalizovala záporná hodnota logaritmickej funkcie maximálnej vierohodnosti penalizovaná počtom parametrov v modeli. Snahou je teda, aby bolo AIC čo najmenšie. Je popísané rovnicou v tvare

$$AIC = -2\log(L) + 2p,$$

kde $\log(L)$ je hodnota logaritmu funkcie maximálnej vierohodnosti a p je počet parametrov v modeli.

6.2 Bayesovo informačné kritérium BIC

Ďalším rozšíreným kritériom na výber modelu je Bayesovo informačné kritérium BIC [41]. Na rozdiel od Akaikeho, Bayesovo informačné kritérium používa na penalizáciu okrem informácie o počte parametrov aj logaritmus dĺžky časového radu, ktorý analyzujeme.

41 Yang, Y. Can the strengths of AIC and BIC be shared? A conflict between model identification and regression estimation. 2005.

BIC je definované rovnicou

$$BIC = -2 \log(L) + p \log(n),$$

pričom n je počet analyzovaných dát a p je počet parametrov v modeli. Taktiež sa pri výbere parametrov modelu snažíme toto kritérium minimalizovať.

6.3 Test stacionarity dát

Predtým ako začneme dáta modelovať pomocou ARMA modelov, musíme overiť, či sú dáta stacionárne. V prípade, že majú jednotkový koreň, namiesto ARMA modelu použijeme model ARIMA.

Najznámejším testom na odhalenie jednotkového koreňa v dátach je **Dickey-Fullerov test** navrhnutý v roku 1984 [28]. Uvažujme AR(1) proces daný vzťahom (4). S parametrom $p = 1$ máme rovnicu popisujúcu model v tvare

$$X_t = \delta + \alpha_1 X_{t-1} + u_t.$$

Pretože výsledky by boli v prípade nestacionarity skreslené, odpočítame z oboch strán X_{t-1} a dostávame rovnicu

$$X_t - X_{t-1} = \delta + \alpha_1 X_{t-1} - X_{t-1} + u_t.$$

Rovnicu môžeme upraviť na tvar

$$X_t - X_{t-1} = \delta + (\alpha_1 - 1)X_{t-1} + u_t,$$

a keď definujeme $(\alpha_1 - 1) = \theta$, dostávame vzťah

$$X_t - X_{t-1} = \delta + \theta X_{t-1} + u_t.$$

Pri Dickey-Fullerovom teste testujeme nulovú hypotézu, že dáta majú jednotkový koreň, čiže sú nestacionárne. Teda

$$H_0: \theta = 0 \quad vs. \quad H_1: \theta < 0.$$

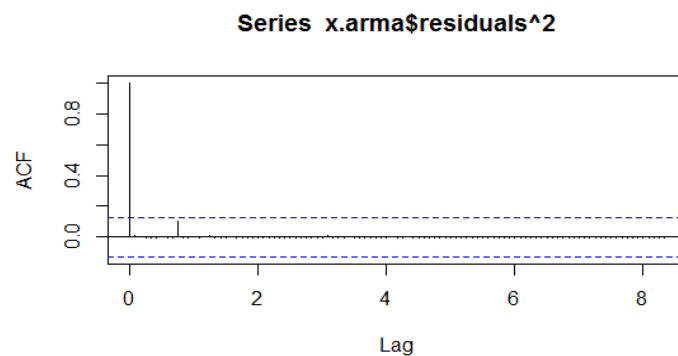
Na vyhodnotenie sa používa t-test významnosti regresných koeficientov s tabelovanými kritickými hodnotami.

28 SAID, S. E., DICKEY, D.A. Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order. 1984.

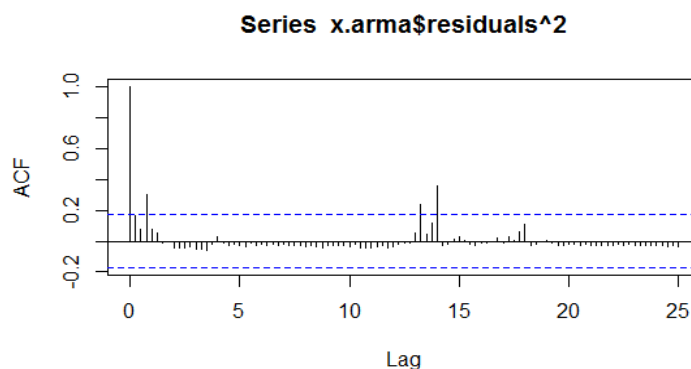
Mnoho makroekonomických premenných vykazuje aj dlhodobý trend, preto je Dickey-Fullerov test rozšírený aj o trendovú funkciu. Takýto test nazývame **Rozšírený Dickey-Fuller test** (*Augmented Dickey-Fuller test*) **ADF test**. Nulová hypotéza zostáva rovnaká. V prípade, že výsledná p-hodnota testu je menšia ako 0,05, nulovú hypotézu zamietame, t.j. prijímame hypotézu, že dáta sú stacionárne. ADF test budeme používať pri analýze dát aj my.

6.4 Test homoskedasticity

Na testovanie homoskedasticity sa používa viacero testov. Na určenie heteroskedasticity procesu sa používajú rezíduá odhadnutého modelu. Na základe korelogramu druhých mocnín rezíduí odhadnutého ARMA modelu môžeme určiť homoskedasticitu alebo heteroskedasticitu procesu. Pri homoskedasticite iba prvá hodnota autokorelačnej funkcie druhých mocnín rezíduí môže byť nenulová. Ako príklad uvádzam grafy z aplikačnej časti tejto diplomovej práce.



Obrázok č. 3: Korelogram druhých mocnín rezíduí ARMA (5,5) modelu vývoja úrokovej miery v Českej republike od roku 1993 do konca roku 2012.



Obrázok č.4: Korelogram druhých mocnín rezíduí ARMA (3,0) modelu vývoja HDP v Belgicku od roku 1980 do roku 2012.

Na obrázku č. 3 vidíme, že len prvá hodnota ACF druhých mocnín rezíduí ARMA modelu presahuje hranice $\pm \frac{2}{\sqrt{n}}$, čo znamená, že je signifikantne odlišná od nuly. V takom prípade je proces homoskedastický. A na obrázku č. 4 vidíme, že nielen prvá, ale viaceré ďalšie hodnoty prekračujú hranice, čiže sú signifikantne väčšie ako 0. V takom prípade je proces heteroskedastický. Grafické zobrazenie autokorelačnej funkcie druhých mocnín rezíduí budeme používať na určenie heteroskedasticity procesu v aplikačnej časti.

Homoskedasticitu je možné zistiť aj pomocou viacerých testov. Veľmi často používaným je **Whitov test** navrhnutý H. Whitom v roku 1980 [38]. Tento test je založený na Lagrangeových multiplikátoroch. Po nájdení vhodného AR (p) modelu v tvare (4)

$$X_t = \delta + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_p X_{t-p} + u_t$$

vypočítame mocniny odhadov rezíduí. Následne zostavíme autoregresný model AR(m) pre mocniny rezíduí

$$\hat{u}_t^2 = \varphi + \varphi_1 \hat{u}_{t-1}^2 + \varphi_2 \hat{u}_{t-2}^2 + \dots + \varphi_m \hat{u}_{t-m}^2 + \varepsilon_t.$$

Nulová hypotéza Whitovho modelu je, že proces je homoskedastický, čiže

$$H_0: \varphi_1 = \varphi_2 = \dots = \varphi_m = 0.$$

Testovacia štatistika je nR^2 , kde n je dĺžka časového radu a R^2 je koeficient determinácie. Pri zisťovaní platnosti nulovej hypotézy sa testovacia štatistika porovnáva s X^2 rozdelením s m stupňami voľnosti. Ak je p-hodnota väčšia ako 0,05, nulovú hypotézu zamietame a predpokladáme, že rezíduá po AR (p) procese sú heteroskedastické a na popis dát by sme mali použiť AR(p) – ARCH(m) alebo AR(p) - GARCH model, prípadne jeho rozšírenia.

6.5 Test normality rezíduí

Normalita rezíduí je dôležitý predpoklad pre interpretáciu odhadu parametrov modelu, ale najmä pre testovanie autokorelácie rezíduí a výpočet štandardných chýb odhadov parametrov. **Jarque-Bera test** navrhnutý C.M. Jarquem a A. K. Berom v roku

38 WHITE, H. A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. 1980.

1980 [20] je založený na súčasnom testovaní šikmosti a špicatosti. Vychádza sa zo skutočnosti, že šikmosť (*skewness*) normálneho rozdelenia je 0 a špicatosť (*kurtosis*) normálneho rozdelenia je 3.

Majme časový rad x_t so strednou hodnotou μ a rozptylom σ^2 . Označme výberovú šikmosť SK a výberovú špicatosť K . Potom je testovacia štatistika Jarque-Berovho testu

$$JB = \frac{n}{6}(SK^2 + \frac{1}{4}(K - 3)^2),$$

pričom n je dĺžka analyzovaného časového radu. Za predpokladu platnosti nulovej hypotézy, ktorá znamená normalitu rezíduí, majú štatistiky SK a K asymptoticky normované normálne rozdelenie $N(0, 1)$ a štatistika JB má rozdelenie χ^2 s dvoma stupňami voľnosti. K zamietnutiu nulovej hypotézy často dochádza aj v prípade, ak rezíduá nemajú konštantný rozptyl, teda nie sú homoskedastické. Aj tento test budeme používať v aplikačnej časti našej práce.

6.6 Testovanie nezávislosti

Pri testovaní nezávislosti analyzovaných dát zisťujeme, či časový rad pozostáva z náhodných hodnôt. **Box-Pierce test** navrhnutý v roku 1970 [6] má ako nulovú hypotézu nezávislosť analyzovaných dát. Box-Piercova testovacia štatistika má tvar

$$Q = n \sum_{k=1}^h r_k^2,$$

kde n je počet pozorovaní, h je maximálny počet hodnôt ACF, ktoré uvažujeme a r_k^2 je mocnina k -tej hodnoty výberovej autokorelačnej funkcie rezíduí. Testovaciu štatistiku Q porovnáваме s χ^2 rozdelením, ktoré má za predpokladu platnosti nulovej hypotézy $(h - m)$ stupňov voľnosti, kde m je počet parametrov odhadovaného modelu. Ak zisťujeme nezávislosť dát, $m = 0$. Tento test sa však často používa aj pri testovaní nezávislosti rezíduí odhadnutého modelu.

20 JARQUE, C.M., BERA, A. K. Efficient tests for normality, homoscedasticity and serial independence of regression residuals.1980.

6 BOX, G. E. P., PIERCE, D. A. Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models. 1970.

Na testovanie nezávislosti sa taktiež veľmi často využíva **Ljung-Boxov test** navrhnutý v roku 1978 [25]. Používa sa Ljung-Boxova Q štatistika na testovanie nulovej hypotézy, že v časovom rade nie je autokorelácia až po rád p . Tento test pracuje na rovnakom princípe ako Box-Pierceov test. Q štatistika sa tiež porovnáva s χ^2 rozdelením s $(h - m)$ stupňami voľnosti a má tvar

$$Q_{LB} = n(n + 2) \sum_{k=1}^p \frac{r_k^2}{n - k}.$$

V aplikačnej časti tejto diplomovej práce používame aj tieto testy nezávislosti.

6.7 BDS test nezávislosti

Odkedy Brock, Dechert a Scheickman v roku 1987 [7] navrhli BDS test, rýchlo sa rozšíril do ekonometrickej praxe. Nulová hypotéza je, že dáta sú generované procesom typu i.i.d. (nezávislé rovnako rozdelené náhodné premenné). BDS test sa používa najmä na rezíduá odhadnutého modelu za účelom testovania zvyšnej závislosti. Zamietnutie nulovej hypotézy znamená, že odhadnutý model nie je dobre špecifikovaný.

Myšlienka tohto testu je nasledovná: Nech x_t je jednorozmerný časový rad generovaný procesom typu i.i.d. Definujme

$$P_A = Pr(|x_t - x_s| < \varepsilon)$$

ako pravdepodobnosť, že dve hodnoty časového radu sú v menšej vzdialenosti ako ε , kde ε je vopred daná presnosť. Ďalej nech

$$P_B = Pr(|x_t - x_s| < \varepsilon \wedge |x_{t-1} - x_{s-1}| < \varepsilon),$$

Pričom P_B je pravdepodobnosť, že súčasné aj predchádzajúce pozorovania sú od seba v menšej vzdialenosti ako ε . Vďaka nezávislosti časového radu, dve vyššie popísané skutočnosti sú v rámci pravdepodobnosti udalosti B nezávislé, a preto $P_B = P_A^2$. Hodnotu týchto pravdepodobností je možné odhadnúť.

25 LJUNG, G.M., BOX, G.E.P. On a Measure of a Lack of Fit in Time Series Models. 1978.

7 BROCK, W.A., DECHERT, W. D., SCHEINKMAN, J. A. A Test for Independence Based on the Correlation Dimension. 1987.

Nulová hypotéza tohto testu je v tvare

$$H_0: P_B - P_A^2 = 0.$$

Na odhadnutie pravdepodobnosti, že dva vektory dĺžky m majú menšiu vzdialenosť ako ε definujeme:

$$c_{m,n}(\varepsilon) = \frac{2}{(n-m+1)(n-m)} \sum_{s=m}^n \sum_{t=s+1}^n \prod_{j=0}^{m-1} I_\varepsilon(x_{s-j}, x_{t-j}),$$

kde

$$I_\varepsilon(x_{s-j}, x_{t-j}) = \begin{cases} 1, & \text{ak } |x_{s-j} - x_{t-j}| < \varepsilon \\ 0 & \text{inak} \end{cases}.$$

Pritom n je veľkosť vzorky dát a m je dimenzia spätného posunutia. V prípade platnosti nulovej hypotézy, teda procesu typu i.i.d., platí

$$E(c_{m,n}(\varepsilon)) = E(c_{1,n}(\varepsilon))^m.$$

Brock [7] dokázal, že BDS štatistika pre dané posunutie m a rádius (presnosť) ε má tvar

$$w_{m,n}(\varepsilon) = \sqrt{n-m+1} \frac{c_{m,n}(\varepsilon) - c_{1,n-m+1}^m(\varepsilon)}{\sigma_{m,n}(\varepsilon)}$$

a asymptotické rozdelenie $N(0,1)$. Konzistentný odhad $\sigma_{m,n}^2(\varepsilon)$ je daný

$$\sigma_{m,n}^2(\varepsilon) = 4(k^m + 2 \sum_{j=1}^{m-1} k^{m-j} c^{2j} + (m-1)^2 c^{2m} - m^2 k c^{2m-2}),$$

kde $c = c_{1,n}(\varepsilon)$,

$$k = k_{1,n}(\varepsilon) = \frac{6}{n(n-1)(n-2)} \sum_{t=1}^n \sum_{s=t+1}^n \sum_{r=s+1}^n h_\varepsilon(x_t, x_s, x_r),$$

$$h_\varepsilon(i, j, k) = \frac{1}{3} [I_\varepsilon(i, j)I_\varepsilon(j, k) + I_\varepsilon(i, k)I_\varepsilon(k, j) + I_\varepsilon(j, i)I_\varepsilon(i, k)].$$

Test je obojstranný, čiže nulová hypotéza procesu typu i.i.d. sa zamieta na hladine významnosti 0,05, ak pre testovaciu štatistiku platí: $|w_{m,n}(\varepsilon)| > 1,96$. Tento test budeme veľmi často využívať v aplikačnej časti diplomovej práce.

7 BROCK, W.A., DECHERT, W. D., SCHEINKMAN, J. A. A Test for Independence Based on the Correlation Dimension. 1987.

II. APLIKAČNÁ ČASŤ

7. Výber a použitie štatistického softvéru

Na analýzu a modelovanie časových radov existuje veľa programov, väčšinou však nie sú voľne dostupné. Aj preto sme si ako program, pomocou ktorého sme modelovali ekonomické a finančné procesy vybrali štatistický softvér R. Je to prostredie a programovací jazyk pre štatistické výpočty a grafiku. Jedná sa o GNU (operačný systém unixového typu), ktorý je podobný programu S vyvinutému v Bellových laboratóriách Johnom Chambersom a jeho kolegami v roku 1976 [43]. Program R môže byť považovaný za implementáciu programu S. Veľkou výhodou oproti iným programom, ako napríklad Matlab, Mathematica, E-view a pod. je, že R je voľne dostupný a zároveň rovnako kvalitný program. Poskytuje širokú škálu štatistických a grafických techník, akými sú lineárne, nelineárne modelovanie, analýza časových dát, klasické štatistické testy, klasifikácia atď. Je veľmi jednoducho rozšíriteľný pomocou inštalácie tzv. balíkov, ktoré sú taktiež voľne dostupné. Jazyk R je často využívaný vo vede a výskume využívajúcim štatistické metódy.

Jednou z výhod programu R je nenáročnosť samotného jazyka, kódu. Ďalej je to ľahkosť, s ktorou sa dajú vytvárať dobre navrhnuté a kvalitné grafy obsahujúce matematické symboly a vzorce. Program R poskytuje efektívne zaobchádzanie s dátami a veľký súdržný a integrovaný súbor čiastkových nástrojov pre analýzu dát, s ktorým nie je zložité pracovať. Preto je často užívaný aj odborníkmi a mnohými svetovými aj slovenskými inštitúciami.

Pri práci sme využívali konkrétne verziu R 2.15.3 a prostredie nazývané R Studio, ktoré je prehľadné a efektívne. Na analýzu časových radov bola potrebná inštalácia niektorých balíkov, ktoré obsahujú funkcie potrebné pre popis dát jednotlivými modelmi a následné predpovedanie. Sú to nasledovné balíky:

43 R Development Core Team. R: A language and environment for statistical computing. 2012.

- tseries - balík na všeobecnú analýzu časových radov, obsahuje ARMA modely, základné GARCH modely a príslušné testy,
- fUnitroot - balík obsahujúci testy stacionarity dát,
- Arfima – balík pre odhad ARFIMA modelov,
- fGarch – balík obsahujúci GARCH modely,
- ruGarch – balík pre využívanie rozšírení triedy GARCH modelov (GJR-GARCH, TGARCH a AVGARCH),
- tsDyn – balík obsahujúci nelineárne modely, akými sú NAAR, STAR, SETAR, LSTAR a príslušné testy,
- forecast – balík pre vytváranie predpovedí.

Na odhad Markov-Switching Autoregresných modelov MSW-AR je v programovacom jazyku R dostupný balík MSwM, ktorý však pracuje s rezíduami lineárnej a nelineárnej regresie. To nie je náš prípad, keďže modelované dáta sú vždy len jeden vektor bez ďalších premenných. Z toho dôvodu sme na analýzu dát pomocou MSW-AR modelu použili program v jazyku Mathematica, ktorý vytvorila Mgr. Lenčuchová, študentka 4. ročníka 3. stupňa študijného programu Aplikovaná matematika na Stavebnej fakulte Slovenskej technickej univerzity v Bratislave.

8. Analyzované dáta

Keďže témou diplomovej práce je modelovanie ekonomických a finančných časových radov, analyzované dáta sú z ekonomického a finančného prostredia. Snažili sme sa o výber historických dát, ktorých časový rad je dostatočne dlhý na analýzu a sú zároveň relevantné z ekonomického hľadiska. Viaceré sú indikátormi trvale udržateľného rozvoja ekonomiky a parametre často zohľadňované pri posudzovaní ekonomiky štátu. Ide o makroekonomické ukazovatele: hrubý domáci produkt (HDP), inflácia, zamestnanosť, čistý export a vývoj úrokovej miery, vývoj menového kurzu a kompozitný indikátor systematického stresu CISS. Dátami sú údaje z viacerých krajín. Zdrojom všetkých analyzovaných dát, okrem údajov CISS, je databáza štatistického úradu Európskej komisie Eurostat [45]. Zdrojom dát CISS je databáza Európskej centrálnej banky Statistical Data Warehouse [47], ktorá tieto hodnoty zverejňuje na svojej internetovej stránke. Časové rady sú rôznej dĺžky podľa dostupnosti údajov v databáze.

8.1 Hrubý domáci produkt HDP

Hrubý domáci produkt HDP je významným indikátorom udržateľného rozvoja. Je to základný makroekonomický ukazovateľ, ktorým meriame celkovú ekonomickú činnosť krajiny. Jeho výška závisí od objemu reprodukčného procesu, ktorý z hľadiska národného hospodárstva predstavuje kolobeh všetkých výrobných faktorov v krajine a prostredníctvom nich vyprodukovaných tovarov a služieb. Je to hodnota všetkých finálnych výrobkov a služieb vyrobených rezidentskými jednotkami štátu za rok. Zodpovedá teda tzv. hrubej pridanej hodnote celej ekonomiky v trhových cenách.

Zdrojové dáta sú sezónne neočistené hodnoty HDP v šiestich vybraných krajinách Európy: Slovensko, Maďarsko, Veľká Británia, Francúzsko, Belgicko a Fínsko. Dáta sú kvartálne a jednotkou je milión Eur od dátumu 1.1.1999 a milión ECU do dátumu 31.12.1998. ECU je skratka pre európsku menovú jednotku (*European currency unit*),

45 European Commission. Statistics Database. [cit.10.03.2013] Dostupné na internete
<http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database>

47 European Central Bank. Statistical data warehouse. [cit. 18.03.2013] Dostupné na internete
<<http://sdw.ecb.europa.eu/>>

ktorá bola košovou menovou jednotkou krajín Európskych spoločenstiev slúžiaca na zúčtovanie medzinárodných operácií do konca roku 1998.

8.2 Čistý export

Export a import patria medzi ukazovatele zahraničného obchodu. Čistý export je rozdiel medzi exportom a importom krajiny. Export je objem vyvezeného tovaru a služieb do zahraničia. Import predstavuje tovary a služby vyrobené v zahraničí, ale spotrebované domácimi spotrebiteľmi. Zdrojové dáta sú kvartálne, sezónne neočistené rozdiely exportu a importu jednotlivých krajín. Objem exportu aj importu v určitom kvartáli je v cenách v predchádzajúcom roku. Jednotkou je milión Eur. Historické údaje sú zo šiestich krajín Európy: Slovensko, Maďarsko, Veľká Británia, Francúzsko, Švajčiarsko a Fínsko.

8.3 Inflácia

V Európskej únii sa inflácia spotrebiteľských cien meria prostredníctvom Harmonizovaného indexu spotrebiteľských cien HICP. Slovo harmonizovaný tu znamená, že všetky členské štáty EÚ dodržiavajú rovnakú metodológiu na vypočítanie indexu. Tento index zaručuje porovnateľnosť medzi krajinami v rámci celej EÚ. Aby bol HICP relevantný, musí čo najpresnejšie odrážať spotrebiteľské správanie domácností v každodennom živote. Aby to bolo možné dosiahnuť, zostavujú národné štatistické úrady v členských štátoch „kôš“ tovarov a služieb, ktorý odráža reprezentatívnu štruktúru spotreby pre všetky typy domácností v danej krajine. Tento kôš zahŕňa ceny bežne kupovaných tovarov a služieb ako je nájom, potraviny, oblečenie, kadernícke služby, no aj menej bežných, ale opakujúcich sa nákupov, ako sú poisťky, opravy a údržba auta či domu. Ceny týchto tovarov a služieb sa zaznamenávajú v pravidelných odstupoch. Nie všetky zložky koša majú rovnakú váhu: potraviny sa nakupujú takmer denne, palivo do auta asi týždenne, návšteva kaderníctva prebehne možno raz za mesiac a poistenie auta sa platí raz za rok. Takže zatiaľ čo je zloženie koša HICP reprezentatívnu vzorkou toho, čo sa nakupuje, je stále potrebné posudzovať, koľko kusov sa nakúpi z každej položky, čiže každá položka musí byť vážená, aby sa stanovila jej dôležitosť v rozpočte domácností. Na aplikovanie vážených hodnôt do koša HICP sa tisícky výrobkov a služieb klasifikujú do viac ako 100 skupín. Pravidelné prieskumy v štruktúre spotreby

domácností stanovujú vážené hodnoty, ktoré sa aplikujú na jednotlivé skupiny. Na účely merania celkovej inflácie sa skupiny položiek v koši zatriedujú do 12 kategórií – napríklad potraviny a nealkoholické nápoje, doprava, rekreácia a kultúra, oblečenie a obuv. Práve súhrnné a vážené zmeny cien v tomto koši vytvárajú index spotrebiteľských cien. HICP je teda priemerné číslo a jednotlivé domácnosti sa môžu stretnúť s vyššími či nižšími mierami inflácie u jednotlivých výrobkov [44].

Zdrojové dáta sú sezónne neočistené hodnoty HICP v rovnakých krajinách Európy ako pri HDP: Slovensko, Maďarsko, Veľká Británia, Francúzsko, Belgicko a Fínsko. Údaje sú mesačné, pričom ide o mesačnú zmenu s indexom 2005=100 zachytávajúcu informáciu o všetkých položkách v spotrebnom koši.

8.4 Zamestnanosť

Ďalším typom dát, ktoré sme si vybrali na analýzu, je vývoj zamestnanosti v štáte. Ide o priemerný evidenčný počet zamestnancov za sledované obdobie vo všetkých odborných činnostiach zaradených do klasifikácie NACE. Údaje zahŕňajú stálych i dočasných zamestnancov, ktorí sú v pracovnom, služobnom, štátnozamestnaneckom alebo v členskom pomere k organizácii bez ohľadu na to, či sú skutočne prítomní v práci alebo nie. Patria sem tiež zamestnanci s kratším pracovným časom. Nepatria sem osoby na materskej dovolenke, osoby vo výkone vojenskej služby, uční a študenti na prevádzkovej praxi [46]. Zdrojové dáta sú kvartálne, sezónne neočistené počty zamestnancov s jednotkou 1 000 osôb v rovnakých krajinách Európy ako pri HDP a inflácii: Slovensko, Maďarsko, Veľká Británia, Francúzsko, Belgicko a Fínsko.

8.5 Úroková miera

Úroková miera predstavuje náklady na požičanie peňazí. Na makroekonomickej úrovni sú kľúčové úrokové sadzby spravidla stanovené centrálnymi bankami ako primárny nástroj menovej politiky s cieľom udržať cenovú stabilitu a kontrolu inflácie. Analyzované

44 Európska komisia. Meranie inflácie v EÚ. [cit. 15.04.2013] Dostupné na internete <http://ec.europa.eu/economy_finance/focuson/inflation/measuring_sk.htm>

46 Štatistický úrad Slovenskej republiky. Metodické vysvetlivky. [cit. 15.04.2013] Dostupné z <<http://portal.statistics.sk/showdoc.do?docid=1648>>

dáta sú úrokové sadzby na peňažnom trhu, tiež známe ako medzibankové sadzby. Tie sú používané bankami pri operáciách medzi sebou. Na peňažnom trhu sú banky schopné obchodovať svoje prebytky a deficity. Pracovali sme s mesačnými údajmi vyjadrenými v percentách v šiestich krajinách sveta: USA, Švédsko, Česko, Poľsko, Maďarsko a Lotyšsko.

8.6 Výmenný kurz

Ďalšou skupinou dát, ktoré sme analyzovali sú výmenné kurzy bývalých národných mien krajín Eurozóny voči mene Euro alebo ECU. Historické údaje sú pre zvolené krajiny až do dátumu zavedenia eura ako štátnej meny. Analyzovali sme výmenné kurzy mien nasledovných šiestich krajín: Slovensko (Koruna), Rakúsko (Šíling), Francúzsko (Frank), Taliansko (Líra), Nemecko (Nemecká marka) a Fínsko (Fínska marka). Dáta sú mesačné hodnoty na konci periódy.

8.7 Kompozitný indikátor systematického stresu CISS

Poslednou analyzovanou skupinou dát je kompozitný indikátor systematického stresu na finančných trhoch Európskej únie [18]. Jeho špecifický štatistický dizajn je tvarovaný podľa štandardných definícií systémového rizika. Hlavná metodologická inovácia CISS je aplikácia základnej teórie portfólia na vytvorenie meradiel finančného stresu na jednotlivých finančných trhoch. Indikátor CISS kladie pomerne veľký dôraz na situácie, v ktorých stres prevládne v niekoľkých segmentoch trhu v rovnakom čase. Zachytáva názor, že finančný stres je systematický a tým je nebezpečný pre ekonomiku ako celok, keď sa finančná nestabilita šíri viacerými segmentmi finančného systému. Analyzovali sme 5 časových radov s týždennou periódou. Prvý časový rad obsahuje hodnoty CISS na trhu s dlhopismi, druhý na akciovom trhu, tretí na peňažnom trhu, štvrtý časový rad popisuje hodnoty CISS na devízovom trhu a piaty na trhu finančných sprostredkovateľov. Na každom trhu CISS zachytáva volatilitu rôznych finančných nástrojov.

18 HOLLÓ, D., KREMER, M., DUCA, M. CISS – A Composite Indicator of Systemic Stress in the Financial System. 2012.

8.8 Základné charakteristiky analyzovaných dát

V tabuľkách 1 - 7 sa nachádzajú základné charakteristiky analyzovaných časových radov pred a po dekompozícii. Skratka PredD znamená pred dekompozíciou časového radu, skratka PoD znamená po dekompozícii. V ľavom hornom rohu tabuľky je názov, skratka skupiny časových radov rovnakého typu, ktorú budeme ďalej používať. V stĺpcoch sú uvedené názvy krajín, ktorých historické dáta analyzujeme. V prvom riadku je časové obdobie, v ktorom boli dáta dostupné. V ďalšom riadku je veľkosť vzorky, čiže dĺžka jednotlivých časových radov. Potom je uvedená minimálna a maximálna hodnota časového radu, stredná hodnota a rozptyl. Tieto hodnoty sú počítané pred dekompozíciou časového radu a aj po odstránení trendu a sezónnej zložky, t.j. pre rezíduá po systematických zložkách.

Tabuľka č. 1: Základné charakteristiky dát hrubého domáceho produktu.

HDP	Slovensko	Maďarsko	Veľká Británia	Francúzsko	Belgicko	Fínsko
Vzorka dát	1993Q1-2012Q4	1995Q1-2012Q5	1980Q1-2012Q4	1980Q1-2012Q4	1980Q1-2012Q4	1970Q1-2012Q4
Veľkosť vzorky	80	72	132	132	132	151
Minimum PredD	2780	8409	87512	119226	5778	21233,4
Maximum PredD	18879	27834	523231	521371	50165	98988
Stredná hodnota PredD	9011	18123	290079	312436	26059	54033
Rozptyl PredD	27234412	40175195	17723521889	14071009050	165595299	533522911
Minimum PoD	-690	-2172	-78	-5999	-2257	133
Maximum PoD	766	1498	63	6051	2246	133
Stredná hodnota PoD	-2,82	-0,49	0,08	-0,59	-14,08	1,01
Rozptyl PoD	72876	409188	899	3874236	685252	312139

Tabuľka č. 2: Základné charakteristiky dát čistého exportu.

Čistý export	Slovensko	Maďarsko	Veľká Británia	Francúzsko	Švajčiarsko	Fínsko
Vzorka dát	1994Q1-2012Q4	1996Q1-2012Q5	1980Q1-2012Q4	1980Q1-2012Q4	1980Q1-2012Q4	1976Q1-2012Q4
Veľkosť vzorky	76	68	132	132	128	148
Minimum PredD	-929	-1207	-17362	-14765	-652	-743
Maximum PredD	1174	2573	5664	9645	13632	4431
Stredná hodnota PredD	-113	250	-4153	-559	3615	1047
Rozptyl PredD	157642	769110	30101330	31962141	14306066	1715278
Minimum PoD	-657	-1169	-10932	-1439	-4863	-1372
Maximum PoD	1647	2724	20608	5460	4041	1367
Stredná hodnota PoD	41,38	115,20	406,22	75,22	-28,04	-7,96
Rozptyl PoD	115380	330293	13997123	1573082	931787	145270

Tabuľka č. 3: Základné charakteristiky dát inflácie.

HICP	Slovensko	Maďarsko	Veľká Británia	Francúzsko	Belgicko	Fínsko
Vzorka dát	1996M1-2013M2	1996M1-2013M2	1996M1-2013M1	1996M1-2013M2	1996M1-2013M2	1996M1-2013M2
Veľkosť vzorky	206	206	205	206	206	206
Minimum PredD	-0,40	-0,60	-0,90	-0,60	-1,90	-0,80
Maximum PredD	3,00	3,00	1,00	0,90	2,00	1,00
Stredná hodnota PredD	0,21	0,28	0,18	0,14	-0,06	0,15
Rozptyl PredD	0,11	0,20	0,13	0,09	0,39	0,12
Minimum PoD	-1,39	-1,31	-0,72	-0,49	-1,79	-0,66
Maximum PoD	4,78	1,34	0,58	0,77	1,62	0,90
Stredná hodnota PoD	0,003	0,006	0,002	0,001	-0,004	-0,001
Rozptyl PoD	0,34	0,13	0,04	0,04	0,23	0,06

Tabuľka č. 4: Základné charakteristiky dát zamestnanosti.

Zamestnanosť	Slovensko	Maďarsko	Veľká Británia	Francúzsko	Belgicko	Fínsko
Vzorka dát	1994Q1-2012Q4	1995Q1-2012Q2	1992Q2-2012Q3	1978Q1-2012Q4	1980Q1-2012Q4	1975Q1-2012Q4
Veľkosť vzorky	76	70	82	140	128	152
Minimum PredD	1988	3863	25118	21994	3599	1948
Maximum PredD	2272	4349	29692	27242	4627	2593
Stredná hodnota PredD	2095	4068	27364	24300	3963	2327
Rozptyl PredD	4651	13917	1968954	2819752	86105	22588
Minimum PoD	-19	-44	-133	-211	-44	-31
Maximum PoD	25	37	147	269	46	40
Stredná hodnota PoD	0,06	0,03	-1,93	0,05	-0,03	0,00
Rozptyl PoD	87	321	2374	1605	331	174

Tabuľka č. 5: Základné charakteristiky dát úrokovej miery.

Úroková miera	Česko	Maďarsko	Poľsko	Švédsko	Lotyšsko	USA
Vzorka dát	1993M1-2013M2	1993M1-2013M2	1995M1-2013M2	1985M1-2004M9	1993Q8-2013M2	1975M1-2006M9
Veľkosť vzorky	237	237	213	237	230	381
Minimum PredD	0,19	4,31	2,46	2,10	0,14	0,98
Maximum PredD	71,51	35,13	28,38	39,34	56,60	19,10
Stredná hodnota PredD	5,61	12,73	10,89	7,82	7,79	6,53
Rozptyl PredD	39,25	57,32	62,41	18,73	120,65	12,70
Minimum PoD	-10,35	-3,81	-4,19	-22,55	-8,60	-4,37
Maximum PoD	50,96	4,65	2,64	4,90	14,46	5,34
Stredná hodnota PoD	0,02	0,04	0,00	-0,11	0,30	0,05
Rozptyl PoD	15,01	0,72	0,55	3,31	4,40	0,81

Tabuľka č. 6: Základné charakteristiky dát menových kurzov.

Kurz Euro\	Koruna	Frank	Lira	Marka	Fínska marka	Šiling
Vzorka dát	1992M2-2007M12	1974M7-1992M3	1974M7-1992M3	1974M7-1992M3	1975M6-2013M2	1974M7-1992M3
Veľkosť vzorky	191	295	295	295	284	295
Minimum PredD	30,21	5,17	775,51	1,85	4,26	13,03
Maximum PredD	45,50	7,10	2291,05	3,14	7,05	22,20
Stredná hodnota PredD	38,87	6,40	1454,40	2,25	5,19	15,90
Rozptyl PredD	11,69	0,30	126668,20	0,11	0,37	5,83
Minimum PoD	-2,14	-0,25	-89,00	-0,08	-0,43	-0,66
Maximum PoD	2,12	0,29	174,22	0,13	0,55	0,50
Stredná hodnota PoD	0,001	-0,004	1,30	0,0002	-0,001	0,00
Rozptyl PoD	0,270	0,002	607,52	0,0003	0,01	0,02

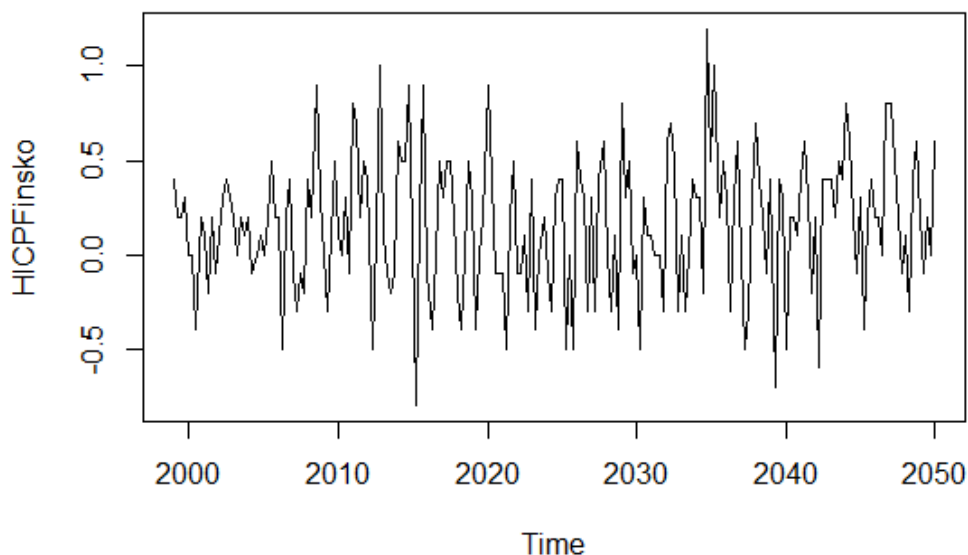
Tabuľka č. 7: Základné charakteristiky dát CISS.

CISS	Trh s dlhopismi	Akciový trh	Peňažný trh	Zahraničný výmenný trh	Trh sprostredkovateľov
Vzorka dát	8.1.1999-5.4.2013	8.1.1999-5.4.2013	8.1.1999-5.4.2013	8.1.1999-5.4.2013	8.1.1999-5.4.2013
Veľkosť vzorky	744	744	744	744	744
Minimum PredD	0,0042	0,0082	0,0058	0,0009	0,0178
Maximum PredD	0,1438	0,2265	0,1447	0,1433	0,2891
Stredná hodnota PredD	0,0526	0,0984	0,0484	0,0502	0,1218
Rozptyl PredD	0,0010	0,0027	0,0007	0,0009	0,0045
Minimum PoD	-0,0651	-0,0909	-0,0455	-0,0798	-0,1043
Maximum PoD	0,0323	0,0697	0,0679	0,0651	0,0754
Stredná hodnota PoD	-0,0004	-0,0002	0,0003	0,0003	-0,0015
Rozptyl PoD	0,0001	0,0003	0,0001	0,0003	0,0003

9. Modelovanie časového radu hodnôt inflácie vo Fínsku

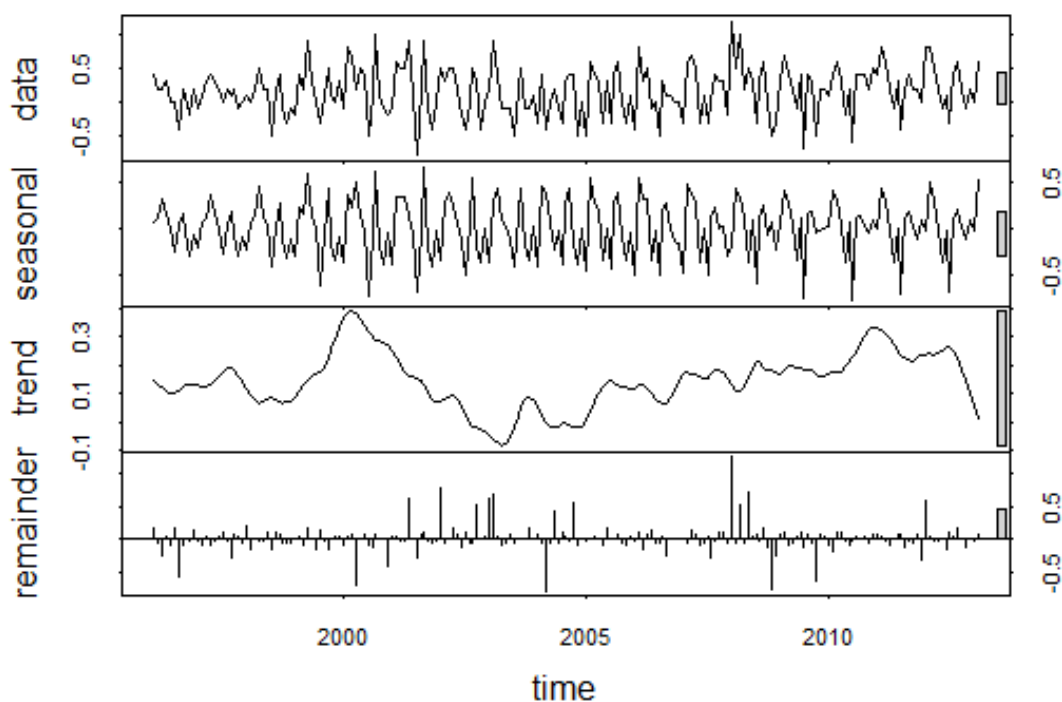
V tejto kapitole diplomovej práce uvedieme názorný postup modelovania jedného časového radu. Presne podľa popisu nižšie sme modelovali aj zvyšných 40 časových radov dát charakterizovaných v kapitole 8. Príslušné príkazy využívané v programe R na modelovanie časového radu vývoja hodnôt inflácie Fínska sú uvedené v Prílohe 1 tejto diplomovej práce. Nachádza sa tam aj popis jednotlivých príkazov. Obdobný kód sme používali pri analýze všetkých dát.

Jednotlivé časové rady sú uložené v textovom súbore ako jeden stĺpcový vektor vrátane hlavičky s názvom časového radu. Po zapnutí programu R, konkrétne prostredia R Studio, ktoré na analýzu využívame, načítame príslušné dáta. Definujeme časový rad s názvom IFinsko s frekvenciou 12, pretože pracujeme s mesačnými dátami. V prípade dát s ročnou periódou je frekvencia 1, v prípade kvartálnej periódy je frekvencia 4 a pod. Ďalej pri definícii časového radu zadáme začiatkový moment, čiže 1. mesiac roku 1996. Grafické zobrazenie tohto časového radu je na nasledovnom obrázku č. 6.



Obrázok č. 6: Vývoj časového radu hodnôt inflácie vo Fínsku.

Na grafe vývoja časového radu vidíme periodické správanie. Urobíme dekompozíciu tohto časového radu. Jej zobrazenie je na obrázku č. 7. Rad má významnú sezónnu zložku a vidíme aj zobrazenie trendu. Pracovať budeme ďalej už len s rezíduami, na obrázku označenými ako remainder. Časový rad rezíduí si označíme IFin.



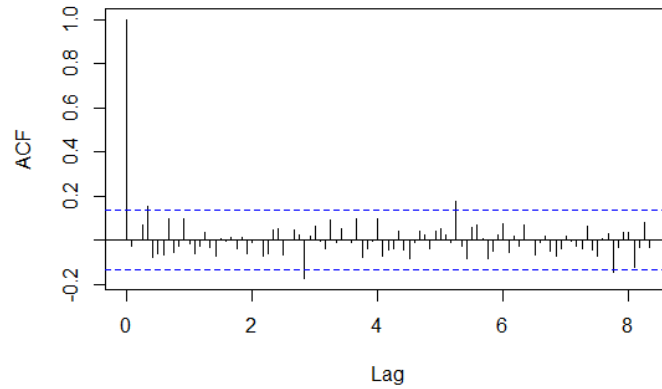
Obrázok č. 7: Dekompozícia časového radu hodnôt inflácie vo Fínsku.

Pred ďalšou prácou s dátami je potrebné nainštalovať si stiahnuté balíky, ktoré budeme potrebovať. Sú popísané v kapitole 7 tejto diplomovej práce. Tieto balíky načítame z knižnice programu R.

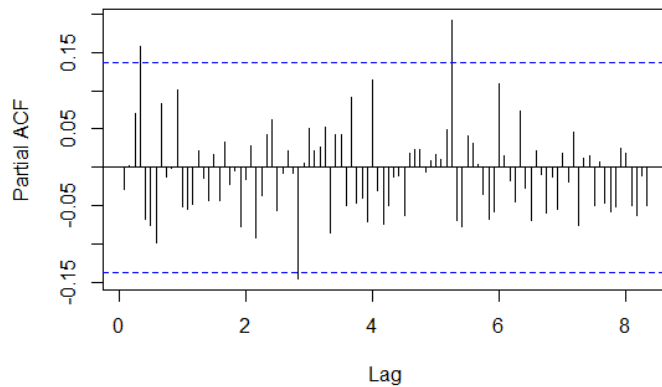
Kvôli posúdeniu predikčných vlastností jednotlivých modelov odoberieme z rezíduí po systematických zložkách posledných h hodnôt a modely odhaduje pomocou ostávajúcich dát dĺžky N .

Teraz by sme mali zistiť, či sú dáta stacionárne alebo obsahujú jednotkový koreň. Použijeme na to ADF test. Výsledná p-hodnota je rovná 0,01, čiže na hladine významnosti 0,05 nulovú hypotézu, že dáta majú jednotkový koreň zamietame. Keďže sú dáta stacionárne, nie je potrebné ich diferencovať. V prípade, že by ADF test nezamietol nulovú hypotézu, museli by sme spraviť prvé diferencie a tie znova otestovať. Ak by už diferencie boli stacionárne, mohli by sme pracovať s nimi, ale ak nie, museli by sme pokračovať v diferencovaní, kým by sme jednotkový koreň neodstránili.

Pozrieme sa na ACF a PACF prvých 100 hodnôt, či nám pomôžu určiť parametre ARMA modelu.



Obrázok č.8: Vývoj autokorelačnej funkcie časového radu hodnôt inflácie Fínska po dekompozícii.



Obrázok č. 8: Vývoj parciálnej autokorelačnej funkcie časového radu hodnôt inflácie vo Fínsku po dekompozícii.

Snažíme sa popísať dáta modelom ARMA. V prípade, že pri výbere určitých parametrov proces nie je stacionárny, program R vypíše hlásenie a je potrebná iná voľba parametrov. Tie vyberieme podľa Akaikeho informačného kritéria a Bayesovho informačného kritéria, ktoré sa snažíme minimalizovať. Napríklad model ARMA (1,0) má $AIC = -43,4$ a $BIC = -33,44$. Model ARMA (0,1) má taktiež $AIC = -43,4$ a $BIC = -33,43$. A model ARMA (1,1) má $AIC = -41,4$ a $BIC = -28,11$. Pridávaním parametrov môžeme tieto hodnoty meniť. Našli sme napríklad model ARMA (0, 4), ktorý má ešte nižšie $AIC = -44,15$, ale vyššie $BIC = -24,22$. Zvyšné modely majú už AIC a BIC vyššie ako model ARMA (1, 0). Preto si na popis inflácie vo Fínsku vyberáme model ARMA (1, 0).

Vykonáme diagnostické testy. Začneme testami nezávislosti rezíduí. V programe R vypočítame jedným príkazom Box-Piercovu a Ljung-Boxovu Q štatistiku. Výsledná p-

hodnota v tomto prípade je 0,9998, čiže je väčšia ako 0,05, čo znamená, že nezamietame nulovú hypotézu nezávislosti rezíduí. Overíme vhodnosť ARMA (1,0) modelu ešte BDS testom, ktorý testuje nezávislosť a rovnaké rozdelenie rezíduí odhadnutého modelu. BDS test s dimenziou $m = 2$ pre tento model vyhodnotí všetky p-hodnoty pre rôzne ε na viac ako 0,05, preto nulovú hypotézu nezávislosti rezíduí a ich rovnakého rozdelenia nezamietame.

Popisné vlastnosti jednotlivých modelov budeme porovnávať pomocou **výberovej štandardnej odchýlky** (*standard deviation*) σ_ε . Výberová štandardná odchýlka sa počíta pomocou vzorca

$$\sigma_\varepsilon = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2},$$

kde N je dĺžka časového radu a μ je stredná hodnota definovaná ako

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i.$$

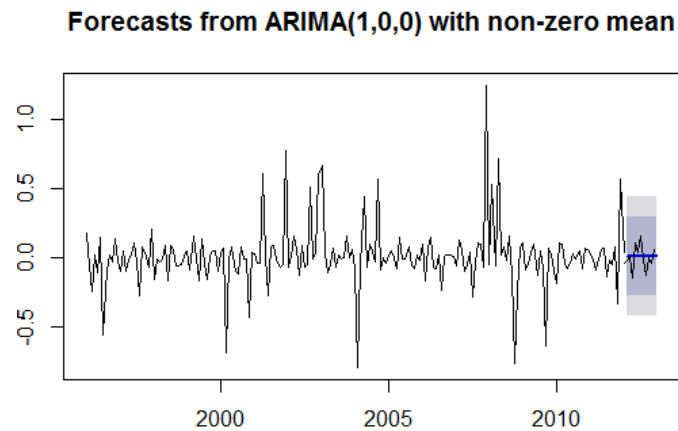
Výberová štandardná odchýlka pri odhadnutom ARMA (1, 0) modeli je $\sigma_\varepsilon = 0,21503$. Pri porovnávaní jednotlivých modelov budeme pozorovať nielen popisné vlastnosti modelu, ale aj predikčné. V prípade časového radu vývoja inflácie vo Fínsku sme z rezíduí po systematických zložkách odobrali posledných 13 hodnôt. Ide o všetky hodnoty v roku 2012 a hodnotu v prvom mesiaci roku 2013. Tie porovnáваме s vytvorenými predpoveďami. Vypočítame teda predpoveď na 13 hodnôt dopredu. Na porovnanie predikčných vlastností budeme používať **odmocninu zo strednej kvadratickej chyby predikcie** (*Root mean square error*) **RMSE** definovanú v tvare

$$RMSE = \sqrt{\frac{1}{h+1} \sum_{t=N+1}^{N+h} (x_t - \hat{x}_t)^2},$$

pričom h je počet predikovaných hodnôt. Samozrejme čím je táto chyba nižšia, tým lepšie predikčné vlastnosti daný model má.

Na obrázku č. 9 vidíme grafické zobrazenie vytvorenej predpovede modelom ARMA (1,0). Táto predpoveď je pre všetkých 13 mesiacov takmer rovnaká, čiže model ARMA (1,0) nemá veľmi dobré predikčné vlastnosti. Je vytvorený aj interval spoľahlivosti

predpovede, do ktorého skutočné hodnoty inflácie v roku 2012 spadajú. Ukazovateľ $RMSE = 0,07845$.

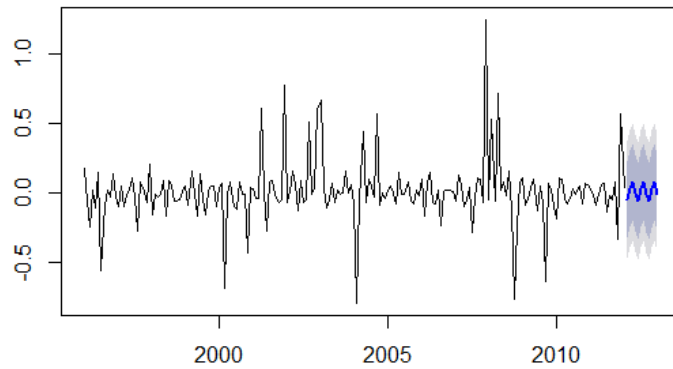


Obrázok č. 9: Predpoveď inflácie na rok 2012 pomocou modelu ARMA (0,1).

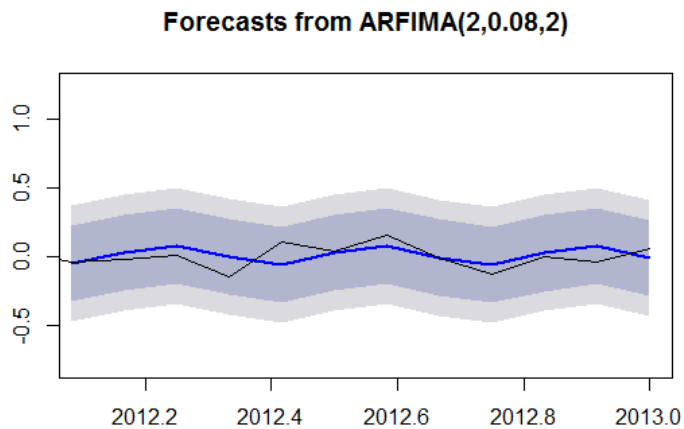
Ďalší model, ktorý použijeme na analýzu dát je model ARFIMA. Odhad parametrov robíme znova pomocou informačných kritérií AIC a BIC. Pri modelovaní časových radov v programe R pomocou modelu ARFIMA, program R sám odhaduje najlepší parameter diferencovania d , pričom vypíše viaceré modely s rôznou hodnotou parametra d . Prvý model v poradí má vždy najnižšie hodnoty AIC a BIC, preto s ním naďalej pracujeme. Ak dáta modelujeme pomocou ARFIMA (1,d,0) dostávame $AIC = -623,300$ a $BIC = -610,008$. Pri použití modelu ARFIMA (0,d,1) dostávame hodnoty $AIC = -623,328$ a $BIC = -610,036$, ktoré sú menšie ako predchádzajúce. Ak použijeme model ARFIMA (1,d,1) dostaneme $AIC = -621,891$ a $BIC = -605,276$. Pridávaním ďalších parametrov sa už v tomto prípade hodnoty AIC a BIC zvyšujú. Preto použijeme model ARFIMA (0,d,1), v ktorom parameter $d = 0,04316$.

Po vykonaní Box-Piercovho a Ljung-Boxovho testu p-hodnota vychádza rovná 0,9876, čo je viac ako 0,05, čiže nezamietame hypotézu o nezávislosti rezíduí. BDS test vychádza pri ARFIMA (0,d,1) modeli so všetkými p-hodnotami väčšími ako 0,05, čiže nezamietame hypotézu o nezávislosti a rovnakom rozdelenia rezíduí ARFIMA (0,d,1) modelu.

Výberová štandardná odchýlka tohto modelu je $\sigma_\varepsilon = 0,21495$. Je takmer rovnaká ako pri modeli ARMA (1,0). Po vytvorení predpovedí pre tento model dostávame chybu predikcie $RMSE = 0,83975$, čo je o niečo viac ako pri ARMA (1,0) modeli.



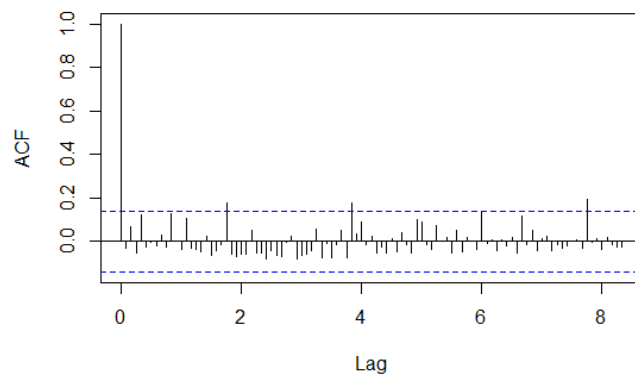
Obrázok č. 10: Predpoveď inflácie na rok 2012 pomocou modelu ARFIMA(0,d,1).



Obrázok č. 11: Porovnanie predpovedi inflácie na rok 2012 pomocou modelu ARFIMA (0,d,1) so skutočnými hodnotami.

Na obrázku č. 10 vidíme ako vyzerá predpoveď inflácie vytvorená modelom ARFIMA (0,d,1) na rok 2012 a aj celý interval spoľahlivosti. Na obrázku č. 11 sú tieto hodnoty približené a porovnané s reálnymi hodnotami inflácie v roku 2012 vyznačenými čiernou farbou. Vidíme, že skutočné hodnoty sú v rámci intervalu spoľahlivosti.

Teraz sa pozrieme na korelogram druhých mocnín rezíduí odhadnutého ARMA (1,0) modelu, aby sme zistili, či je proces homoskedastický.



Obrázok č. 12: Korelogram druhých mocnín rezíduí modelu ARMA (1,0).

Z korelogramu na obrázku č. 12 vidíme, že okrem prvej hodnoty aj niektoré iné presahujú hranicu, čo znamená, že sú štatisticky významne väčšie ako 0. Mocniny rezíduí sú korelované, čiže proces je heteroskedastický. Dáta budeme modelovať pomocou ARCH a GARCH modelu a jeho rozšírení.

GARCH model a jeho rozšírenia modelujeme v programe R pomocou funkcie `garchFit`. Výber parametrov modelu robíme znova podľa AIC a BIC kritérií. GARCH model s najnižšími hodnotami kritérií, aké sa nám podarilo nájsť je ARMA (3,1) + GARCH (1,1) model. V R programe je možné pri všetkých GARCH modeloch vybrať rozdelenie chybovej zložky, ktoré bude model využívať. My sme využívali hlavne zošikmené normálne rozdelenie alebo zošikmené Studentovo rozdelenie. V tomto prípade to bolo zošikmené Studentovo rozdelenie. Pri použití BDS testu dostávame všetky p-hodnoty vyššie ako 0,05, preto nezamietame hypotézu nezávislosti a rovnakého rozdelenia rezíduí tohto modelu.

Na modelovanie vybraných časových radov používame aj rozšírenia triedy GARCH modelov, konkrétne TGARCH, GJRGARCH, AVGARCH. Modelovanie prebieha rovnakým spôsobom ako u štandardného GARCH modelu. Taktiež si môžeme vybrať rozdelenie, ktoré bude použité a snažíme sa minimalizovať výsledné AIC a BIC. Jediným rozdielom je, že pri rozšíreniach modelu GARCH volíme parametre ARMA (p,q) + APARCH (m, n). Najnižšie AIC a BIC sme dostali u všetkých rozšírení GARCH modelu pri výbere parametrov ARMA (3,1) +APARCH (1,1). Výberové štandardné odchýlky rezíduí a RMSE pri vytváraní predpovedí sú v tabuľke č. 8.

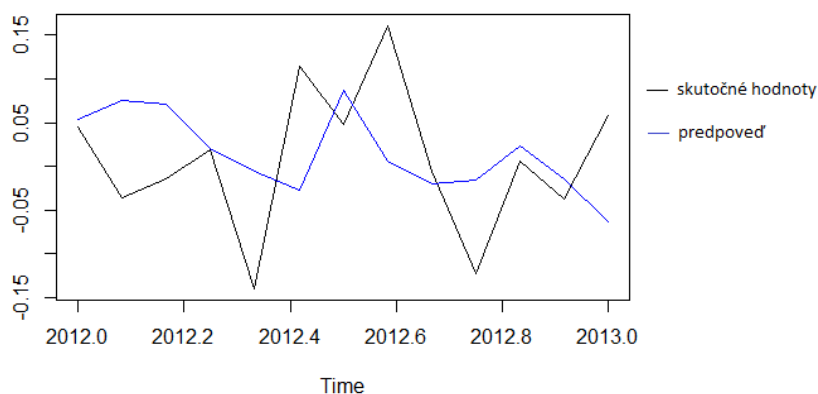
Tabuľka č. 8: Štandardná odchýlka a RMSE modelov triedy GARCH použitých na časový rad hodnôt inflácie vo Fínsku.

	σ_ε	RMSE
GARCH	0,285631	0,07743284
TGARCH	0,285175	0,08068049
GJRGARCH	0,285631	0,07933289
AVGARCH	0,285727	0,07713129

V tabuľke č. 8 vidíme, že výberové štandardné odchýlky rezíduí modelov triedy GARCH použitých v našom prípade vychádzajú veľmi podobne. Najnižšia hodnota σ_ε je u modelu TGARCH. Všetky sú však vyššie ako pri modeloch ARMA(1,0)

a ARFIMA (0,d,1). Najlepšie predikčné vlastnosti hodnotené ukazovateľom RMSE má model AVGARCH.

Na analýzu časového radu hodnôt inflácie vo Fínsku využijeme aj nelineárny aditívny autoregresný model NAAR. V programe R volíme ako parametre modelu maximálny rád m (*embedding dimension*) a časové omeškanie d . Dostupná je pri tomto modeli už len hodnota AIC, ktorej minimalizovaním volíme vhodné parametre. Najnižšie AIC je v prípade voľby $m = 6$ a $d = 6$, konkrétne $AIC = -593$. Pri nižších a aj vyšších hodnotách parametrov m a d je AIC vyššie. Nezávislosť a identickosť rozdelenia rezíduí tohto modelu overíme BDS testom. Všetky p-hodnoty sú vyššie ako 0,05, čiže nulovú hypotézu nezamietame. Štandardná odchýlka v tomto prípade vychádza $\sigma_\varepsilon = 0,1990653$, čo je zatiaľ najnižšia hodnota štandardnej odchýlky, ktorú sme pri tomto časovom rade dostali. Predikcia na 13 mesiacov modelom NAAR je zobrazená na obrázku č. 13. Reálne hodnoty sú znázornené čiernou farbou a predpovedané hodnoty modrou. Pri týchto predpovediach je $RMSE = 0,08925018$.



Obrázok č. 13: Porovnanie predpovede a skutočných hodnôt inflácie vo Fínsku v roku 2012 a v prvom mesiaci roku 2013.

Dáta sme analyzovali aj pomocou viacrežimových modelov. Pri modelovaní prahových autoregresívnych modelov, či už SETAR, STAR alebo LSTAR je potrebné najprv vykonať test linearity. Pri používaní modelov STAR a LSTAR sa tento test v R robí automaticky. Pri každom zadaní parametrov modelu program najprv vyhodnotí, či nie je lepšie použiť lineárny model. Nulová hypotéza testu je, že na modelovanie je lepšie použiť lineárny model. Ak vyjde p-hodnota menšia ako 0,05, nulovú hypotézu zamietame a používame príslušný viacrežimový model. Ak program pri určitom výbere parametrov modelu STAR alebo LSTAR vyhodnotí, že nie sú potrebné viaceré režimy, vypíše

upozornenie, aby sme časový rad modelovali pomocou lineárneho modelu. Často však je možné nájsť vhodné parametre týchto modelov, kedy sú viaceré režimy výhodnejšie ako len jeden.

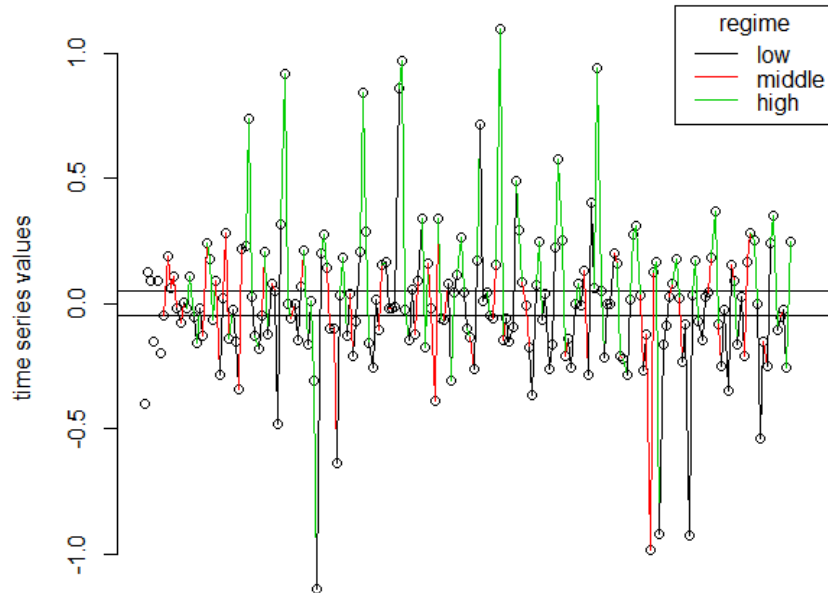
Pri analýze časového radu pomocou SETAR modelu sa tento test lineariry nedeje automaticky. Funkcia programu R s názvom `setarTest` (`data, test="1vs"`) vykonáva dva testy. Prvý test je 1 režim verzus 2 režimy a druhý test 1 režim verzus 3 režimy. Nulové hypotézy týchto testov sú, že jeden režim je lepší ako 2 v prípade prvého testu a 3 v prípade druhého testu. Funkcia vypíše p-hodnoty oboch testov pre zvolené parametre modelu. Ak je p-hodnota menšia ako 0,05, zamietame nulovú hypotézu, čiže môžeme použiť SETAR model. Následne, aby sme zistili, či je lepšie mať dva režimy alebo tri, použijeme funkciu `setarTest` (`data, test="2vs3"`) s nulovou hypotézou, že 2 režimy sú pre daný časový rad postačujúce (pri určitých parametroch modelu SETAR).

Parametre prahových modelov SETAR, STAR aj LSTAR, ktoré meníme a pomocou nich vyberáme model, sú maximálny rád m , časové omeškanie $thDelay$ a maximálny počet režimov $noRegimes$ alebo počet prahov $nthresh$. Omeškanie $thDelay$ musí byť menšie alebo rovné ako maximálny rád m .

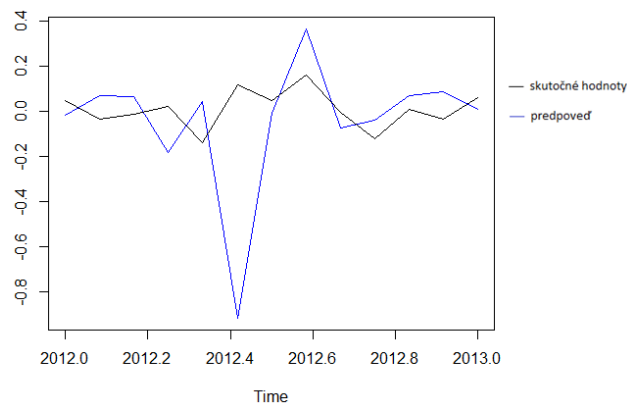
Pri používaní modelu SETAR, program R vyberá najvýhodnejšie hodnoty prahu alebo prahov pre zvolené parametre modelu. Porovnávaním kritéria AIC sme zvolili model SETAR s parametrami $m = 6$, $thDelay = 6$, $nthresh = 2$. Pre takýto model vyšli obe p-hodnoty testu lineariry menšie ako 0,05, teda pri zvolených parametroch jeden režim nestačí a je potrebné pridať ďalšie režimy. Druhý test vypísal p-hodnotu tiež nižšiu ako 0,05, teda pri zvolených parametroch dva režimy nestačia a je potrebné pridať tretí režim. Hodnoty prahov sú $th1 = -0.04876$ a $th2 = 0.04947$. Graf zobrazujúci rozloženie časového radu do jednotlivých režimov ja na obrázku č. 14. Rozdelenie hodnôt do jednotlivých režimov je relatívne rovnomerné. Percentuálny podiel hodnôt v dolnom režime je 39,7%, v strednom je 23,12% a v hornom režime je 37,19% hodnôt časového radu.

BDS test nezávislosti rezíduí tohto SETAR modelu má všetky p-hodnoty väčšie ako 0,05, čiže nezamietame nulovú hypotézu, že reziduá sú nezávislé a rovnako rozdelené. Výberová štandardná odchýlka vybraného SETAR modelu je $\sigma_\varepsilon = 2439639$.

Predpovede vytvorené pomocou zvoleného SETAR modelu sú zobrazené na obrázku č. 15. Ukazovateľ $RMSE = 0,2970628$, čo je dosť veľa oproti $RMSE$ iných modelov, ale je to hlavne kvôli predpovedanej hodnote v 4. mesiaci roku 2012, ktorá je extrémne nízka.



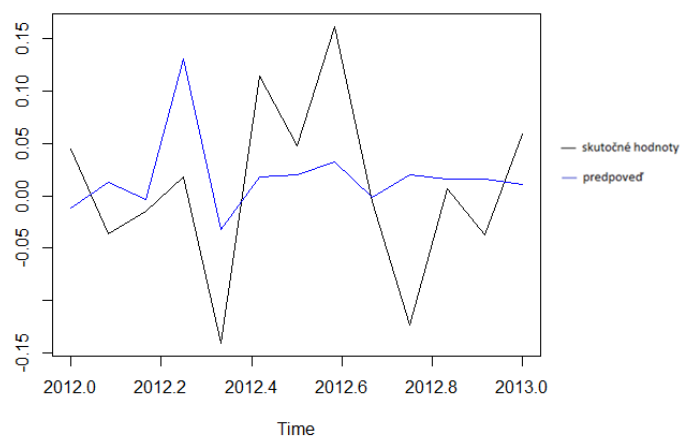
Obrázok č. 14: Graf rozdelenia hodnôt časového radu inflácie vo Fínsku do 3 režimov pomocou modelu SETAR.



Obrázok č. 15: Porovnanie predpovedí vykonaných pomocou zvoleného SETAR modelu a skutočných hodnôt inflácie v roku 2012 a januári 2013.

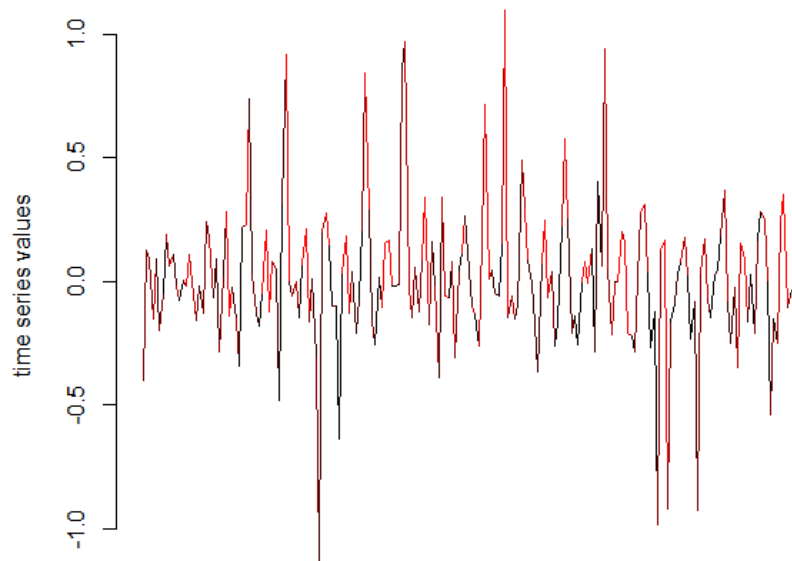
Model STAR sme minimalizovaním AIC kritéria zvolili s parametrami $m = 5$, $thDelay = 3$, $noRegimes = 3$. Program R automaticky robí test linearity, ktorý v tomto prípade vypísal p-hodnotu menšiu ako 0,05, čiže môžeme použiť viacerežimový model STAR. Program najprv vytvorí model STAR s dvoma režimami a následne testuje prídanie ďalšieho režimu. V tomto prípade prídanie 3. režimu program zamietá a vytvoril

dvojrežimový STAR model. Prahovú hodnotu vyberá taktiež automaticky. Výsledný model má hodnotu parametra vyhladenia $\gamma = 10,0002$ a prahová hodnota je $th = 0,1790062$. BDS test nezamieta nulovú hypotézu o nezávislosti a identickosti rozdelenia rezíduí tohto modelu. Výberová štandardná odchýlka vychádza $\sigma_\varepsilon = 0,244178$. Pri vytváraní predpovedí na rok 2012 a január 2013 pomocou zvoleného modelu STAR dostávame $RMSE = 0,07675018$, čo je doteraz najnižšia chyba predpovede, ktorú sme pre časový rad inflácie vo Fínsku dostali. Pozrieme sa na grafické porovnanie predpovedaných a skutočných hodnôt zobrazené na obrázku č. 16. Môžeme povedať, že je to doteraz asi najlepšia predpoveď.



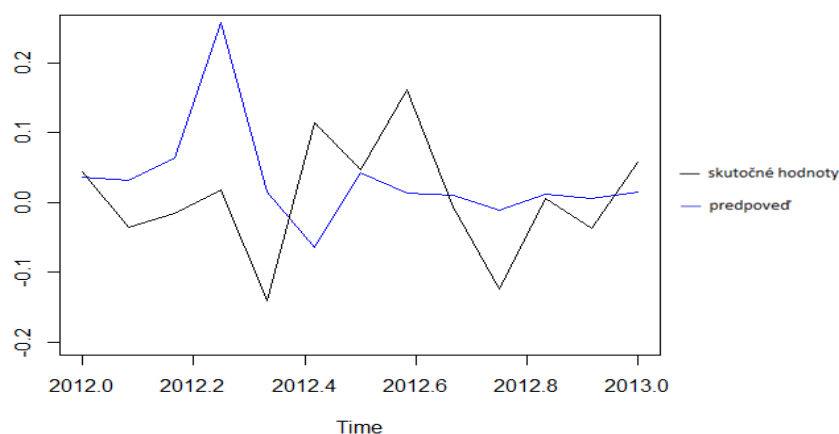
Obrázok č. 16: Porovnanie predpovede vytvorenej modelom STAR so skutočnými hodnotami časového radu inflácie vo Fínsku.

Analýza časových radov pomocou modelu LSTAR je obdobná ako modelom STAR. Minimalizovaním Akaikeho informačného kritéria sme vybrali LSTAR model s maximálnym rádom $m = 6$ a časovým omeškaním $thDelay = 5$. Program R pri modelovaní LSTAR modelu robí automaticky test linearity, ktorý v tomto prípade s p-hodnotou rovnou 0,00792 zamieta nulovú hypotézu, že nie je potrebné pridať ďalší režim, čiže že by stačil lineárny model. Optimálna prahová hodnota pre zvolený LSTAR model je $th = 0,06483$ a parameter vyhladenia $\gamma = 105,4$. BDS test rezíduí vybraného modelu nezamieta nulovú hypotézu nezávislosti a identickosti rozdelenia rezíduí. Grafické zobrazenie rozdelenia hodnôt do jednotlivých režimov je znázornené na obrázku č. 17. Čierna farba popisuje hodnoty v dolnom režime a červená farba znázorňuje hodnoty v hornom režime.



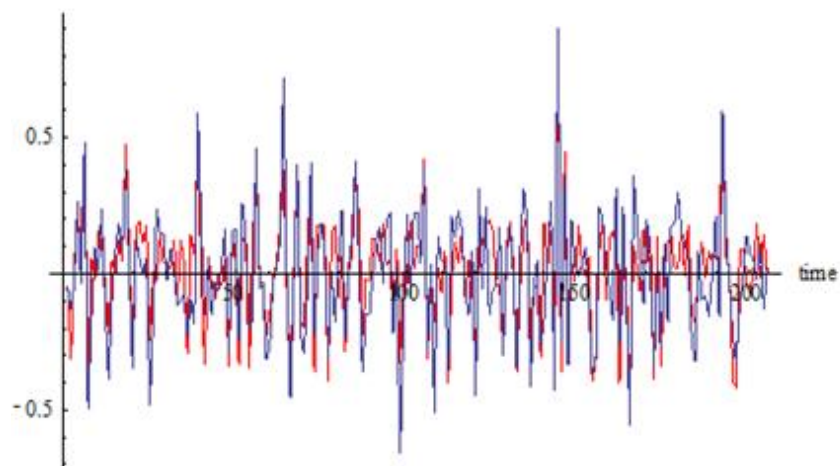
Obrázok č. 17: Graf rozdelenia hodnôt časového radu inflácie vo Fínsku do dvoch režimov pomocou modelu LSTAR.

Výberová štandardná odchýlka vychádza $\sigma_\varepsilon = 0,2553267$. Pri všetkých nelineárnych prahových modeloch časového radu inflácie vo Fínsku vychádza pomerne vysoká oproti modelu ARMA (0,1) alebo ARFIMA (0,d,1). Predikčné vlastnosti tohto modelu hodnotíme ukazovateľom RMSE, ktorý vychádza rovný 0,1077556, čo je pomerne vysoké číslo oproti zvyšným použitým modelom. Grafické porovnanie predpovedí s reálnymi hodnotami je na obrázku č. 18. Z grafu vidíme, že predpovedaná hodnota na druhý mesiac je extrémne vysoká. Aj preto vychádza vyššia chyba predikcie $RMSE = 0,1077556$.



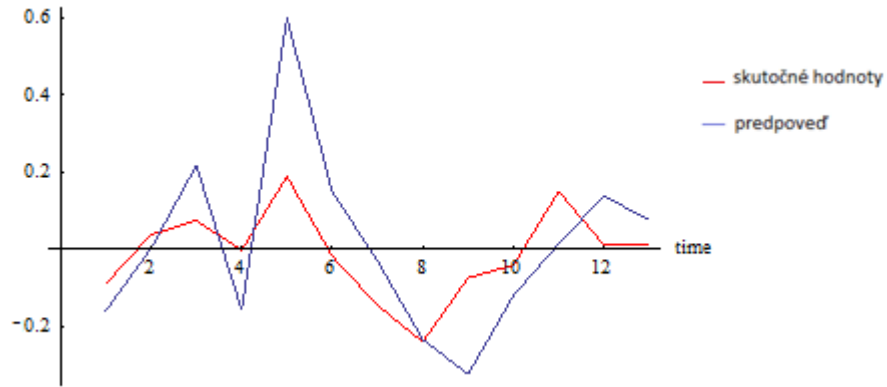
Obrázok č. 18: Porovnanie predpovede vytvorenej modelom LSTAR so skutočnými hodnotami časového radu inflácie vo Fínsku.

Posledný model, ktorý sme používali na analýzu časových radov je Markov-Switching Autoregresný model MSW-AR. Keďže sme nenašli vhodný balík v prostredí R na modelovanie časových radov pomocou MSW-AR modelov, použili sme program v jazyku Mathematica, ktorý vytvorila Mgr. Lenčuchová, študentka 4. ročníka 3. stupňa študijného programu Aplikovaná matematika na Stavebnej fakulte Slovenskej technickej univerzity v Bratislave. Tento program obsahuje okrem odhadov parametrov 2- aj 3-režimových MSW-AR modelov aj testy linearity, ostávajúcej nelinearity, autokorelácie rezíduí vo vnútri režimov aj medzi režimami a BDS test. Program sme použili na časový rad rezíduí po dekompozícii vykonanej v programe R. Výber parametrov modelu sme robili minimalizovaním Bayesovho informačného kritéria BIC pre tie modely, ktoré prešli všetkými testami. Takto sme zvolili trojrežimový model MSW-AR s parametrom $p = 1$. Výsledná výberová štandardná odchýlka je najnižšia, akú sa nám pri modelovaní časového radu inflácie vo Fínsku podarilo dosiahnuť, a to $\sigma_\varepsilon = 0,147553$. Oproti zvyšným modelom aplikovaným na popis spomínaného časového radu je o poznanie nižšia.



Obrázok č. 19: Porovnanie modelu MSW-AR (červená čiara) so skutočnými hodnotami časového radu inflácie vo Fínsku (modrá čiara).

Predikčné vlastnosti hodnotené RMSE sú však porovnateľné s predikčnými vlastnosťami ostatných modelov, nakoľko pri modeli MSW-AR vychádza $RMSE = 0,0814605$. Na obrázku č. 20 vidíme, že vývoj predpovedaných hodnôt vytvorených pomocou modelu MSW-AR zobrazený modrou farbou je relatívne podobný ako vývoj skutočných hodnôt časového radu, aj keď predpovede obsahujú väčšie výkyvy.



Obrázok č. 20: Porovnanie predpovede vytvorenej modelom MSW-AR so skutočnými hodnotami časového radu inflácie vo Fínsku.

BDS testy všetkých modelov potvrdili nezávislosť rezíduí. Najnižšiu štandardnú odchýlku mal model MSW-AR a po ňom NAAR model. Najnižšiu chybu predikcie RMSE sme dosiahli pomocou modelu STAR, avšak podľa grafického zobrazenia najlepšie predikcie sme vytvorili pomocou modelu MSW-AR. Preto za najvhodnejší model pre popis a predikciu časového radu vývoja inflácie vo Fínsku považujeme práve MSW-AR model.

10. Hodnotenie výsledkov modelovania časových radov

Okrem časového radu hodnôt inflácie vo Fínsku sme modelovali aj ďalších 40 časových radov dát popísaných v kapitole 8 tejto práce. Modelovanie prebiehalo obdobným spôsobom ako sme popísali na príklade v predchádzajúcej kapitole. Hlavným kritériom hodnotenia popisných vlastností je výberová štandardná odchýlka rezíduí modelu σ_ε . Lepší je model s menšou σ_ε . Kritériom hodnotenia predikčných vlastností modelu je predpovedná chyba RMSE. Hodnoty týchto ukazovateľov sme pre všetky modely a všetky analyzované dáta zapisovali do tabuliek v Exceli. Pomocou podmieneného formátovania sme vytvorili farebnú škálu v rámci každého analyzovaného časového radu zobrazujúcu najnižšie hodnoty oboch ukazovateľov zelenou farbou, priemerné hodnoty žltou farbou a najvyššie hodnoty červenou farbou.

V prípade, že test linearity nezamietol nulovú hypotézu linearity oproti nelinearite zvoleného typu pre všetky možnosti zvolených parametrov, daný model sme nepoužili a v tabuľke výsledkov sme zapísali NA. Podobne, ak pri teste homoskedasticity nebola zamietnutá nulová hypotéza, modely triedy GARCH sme nepoužili.

Ako prvý sme vždy odhadli model ARMA, ale v prípade, že test stacionarity časového radu potvrdil jednotkový koreň AR-polynómu, a teda nestacionaritu časového radu, použili sme namiesto modelu ARMA model ARIMA. V tabuľke výsledkov sú prípady, kedy sme aplikovali model ARIMA označené hviezdíčkou.

10.1 Výsledky modelovania hrubého domáceho produktu HDP

V tabuľke č. 9 vidíme výsledné hodnoty výberovej štandardnej odchýlky rezíduí σ_ε modelov aplikovaných na časové rady vývoja HDP v šiestich krajinách Európy. Môžeme pozorovať, že najnižšiu hodnotu σ_ε má pre všetky analyzované krajiny model MSW-AR, a to o poznanie nižšiu ako pri zvyšných modeloch. Z ostatných tried modelov najmenšie hodnoty σ_ε majú prahové viacrežimové modely a najvyššie hodnoty σ_ε pozorujeme pri triede GARCH modelov.

Tabuľka č. 9: Výsledné hodnoty štandardnej odchýlky rezíduí modelov časových radov HDP.

Model\Krajina	Veľká					
	Slovensko	Maďarsko	Británia	Francúzsko	Belgicko	Fínsko
ARMA\ARIMA	206,79*	526,12	34,35	1756,80	425,28	439,11
ARFIMA	219,79	557,37	33,99	1859,77	419,84	471,88
GARCH	285,55	646,71	35,37	1917,26	449,54	490,48
TGARCH	286,91	627,82	35,38	2043,20	450,09	502,93
GJRGARCH	284,53	644,30	35,16	1971,58	457,46	490,79
AVGARCH	276,98	652,70	35,58	1961,42	468,03	498,38
NAAR	95,27	517,65	34,08	1839,04	393,09	545,39
SETAR	169,09	389,91	32,24	1679,95	404,02	351,22
STAR	186,19	447,23	NA	1635,20	373,00	376,28
LSTAR	174,83	435,05	31,76	1616,80	375,58	402,55
MSW-AR	56,05	164,38	14,87	1473,38	185,53	246,40

Porovnanie predikčných vlastností pomocou predpovednej chyby RMSE je v tabuľke č. 10. Hodnotenie už nie je také jednoznačné. MSW-AR modely vykazujú RMSE najnižšiu v troch prípadoch. Pri modelovaní časového radu HDP Slovenska vidíme, že namiesto modelu ARMA sme kvôli odstráneniu jednotkového koreňa použili model ARIMA. A práve tento model pri predikcii vykázal najnižšiu hodnotu RMSE. Ale napríklad pri modelovaní časového radu HDP Fínska, model ARMA má najvyššiu hodnotu RMSE. Pri iných krajinách najvyššiu hodnotu RMSE majú rôzne nelineárne modely.

Tabuľka č. 10: Výsledné hodnoty RMSE modelov časových radov HDP.

Model\Krajina	Veľká					
	Slovensko	Maďarsko	Británia	Francúzsko	Belgicko	Fínsko
ARMA\ARIMA	83,22*	974,25	47,47	1044,31	850,26	491,02
ARFIMA	93,06	971,50	47,44	1077,97	930,66	488,03
GARCH	96,78	956,85	46,67	1030,76	944,45	464,80
TGARCH	96,78	958,80	46,65	1033,44	946,53	464,57
GJRGARCH	96,77	957,11	46,88	1023,29	944,62	455,05
AVGARCH	97,27	950,39	47,09	1023,78	943,79	451,08
NAAR	415,73	966,49	47,73	1011,58	878,21	477,69
SETAR	130,15	988,57	45,59	897,51	895,54	442,66
STAR	100,98	1120,63	NA	1024,27	1601,16	466,19
LSTAR	115,26	955,14	47,73	1014,35	774,41	456,99
MSW-AR	119,66	547,86	24,57	1505,72	851,22	313,29

10.2 Výsledky modelovania čistého exportu

Výsledné hodnoty výberovej štandardnej odchýlky rezíduí σ_ε modelov použitých na analýzu časových radov čistého exportu krajín Európy sú zobrazené v tabuľke č. 11. Vidíme, že pri každej krajine najnižšiu hodnotu σ_ε má model MSW-AR, tak ako to bolo pri modelovaní HDP. Najvyššie hodnoty znova dosahujú modely triedy GARCH. Táto podobnosť je logická, nakoľko aj pomocou čistého exportu sa počíta hodnota HDP. Viacrežimové prahové modely vykazujú pomerne nízku výberovú štandardnú odchýlku rezíduí pri všetkých časových radov čistého exportu vybraných krajín Európy.

Tabuľka č. 11: Výsledné hodnoty štandardnej odchýlky rezíduí modelov

časových radov čistého exportu vybraných krajín Európy.

Model\Krajina	Veľká					
	Slovensko	Maďarsko	Británia	Francúzsko	Švajčiarsko	Fínsko
ARMA\ARIMA	250,02*	501,08*	3377,83	1534,49	850,90	643,55
ARFIMA	251,76	504,78	3328,96	1537,37	847,48	650,34
GARCH	275,70	503,56	3472,38	1739,82	952,38	681,24
TGARCH	276,50	515,91	3606,70	1767,97	961,75	693,73
GJRGARCH	286,29	475,39	3657,21	1730,63	961,75	693,57
AVGARCH	281,52	505,42	3568,35	1723,55	962,73	697,18
NAAR	236,98	457,83	3371,00	1702,16	888,80	552,67
SETAR	238,72	390,43	3219,85	1545,29	794,33	626,53
STAR	222,06	432,72	3081,10	1443,22	796,87	574,04
LSTAR	224,06	420,99	2721,34	1674,91	884,23	523,01
MSW-AR	64,21	102,10	1256,44	702,60	402,20	162,14

Tabuľka č. 12: Výsledné hodnoty RMSE modelov časových radov čistého exportu krajín Európy.

Model\Krajina	Veľká					
	Slovensko	Maďarsko	Británia	Francúzsko	Švajčiarsko	Fínsko
ARMA\ARIMA	222,89*	798,68*	4822,84	1560,15	2670,66	1634,61
ARFIMA	319,54	698,33	4786,12	1562,63	2428,93	1572,85
GARCH	214,98	769,53	4920,16	1662,60	2604,71	1551,99
TGARCH	209,47	730,73	4919,30	1661,63	2604,83	1527,98
GJRGARCH	206,21	713,47	4918,46	1661,63	2604,63	1536,43
AVGARCH	224,24	727,54	4884,46	1663,67	2552,54	1530,32
NAAR	221,58	727,33	4742,00	1651,51	2496,98	1595,54
SETAR	274,21	1064,15	4861,36	1932,99	2380,36	1521,03
STAR	282,96	630,24	4650,80	1629,08	2270,94	1413,31
LSTAR	297,06	722,72	4779,86	1650,03	2476,64	1526,18
MSW-AR	141,26	244,34	3667,65	943,83	1202,91	267,37

Výsledné hodnoty RMSE predikcií použitých modelov na časové rady čistého exportu vidíme v tabuľke č. 12. Jednoznačne môžeme povedať, že najnižšie RMSE majú predikcie vytvorené pomocou modelu MSW-AR. Ďalej vidíme, že SETAR model vytvoril predikcie časových radov čistého exportu štyroch krajín s vysokou chybou RMSE. Oproti tomu model STAR vykazuje vo viacerých prípadoch nízke hodnoty RMSE.

V prípade modelovania čistého exportu majú najnižšie hodnoty σ_ε aj RMSE modely MSW-AR. Majú teda najlepšie popisné aj predikčné vlastnosti spomedzi použitých modelov aplikovaných na časové rady čistého exportu.

10.3 Výsledky modelovania inflácie

Modelovanie časového radu inflácie vo Fínsku sme popísali v kapitole č. 9. Okrem toho sme modelovali vývoj inflácie aj v ďalších piatich krajinách Európy. V tabuľke č. 13. vidíme výsledné hodnoty výberovej štandardnej odchýlky rezíduí σ_ε použitých modelov. Znova pozorujeme, že najnižšiu hodnotu σ_ε majú Markov-Switching autoregresné modely. Nízke hodnoty dostávame aj použitím ARMA modelov.

Tabuľka č. 13: Výsledné hodnoty štandardnej odchýlky rezíduí modelov časových radov hodnôt inflácie vo vybraných krajinách Európy.

Model\Krajina	Veľká					
	Slovensko	Maďarsko	Británia	Francúzsko	Belgicko	Fínsko
ARMA\ARIMA	0,490	0,304	0,154	0,160	0,321	0,215
ARFIMA	0,548	0,362	0,187	0,185	0,390	0,215
GARCH	0,507	0,325	0,167	0,170	0,381	0,286
TGARCH	0,529	0,325	0,166	0,170	0,382	0,285
GJRGARCH	0,509	0,325	0,167	0,170	0,379	0,286
AVGARCH	0,511	0,325	0,166	0,169	0,381	0,286
NAAR	0,556	0,306	0,191	0,196	0,381	0,199
SETAR	0,536	0,319	0,184	0,191	0,301	0,244
STAR	0,432	0,321	0,154	0,301	0,301	0,244
LSTAR	0,443	0,307	0,171	0,176	0,404	0,255
MSW-AR	0,271	0,277	0,142	0,145	0,272	0,148

Predikčné vlastnosti modelov posudzujeme podľa ukazovateľa RMSE. Výsledné hodnoty sú zobrazené v tabuľke č. 14. Nevieme jednoznačne určiť jednu triedu modelov, ktorá by mala najlepšie predikčné vlastnosti v prípade analyzovaných časových radov hodnôt inflácie. MSW-AR modely majú najnižšie hodnoty RMSE u časových radov

vývoja inflácie v Maďarsku, vo Veľkej Británii a vo Francúzsku. Pri modelovaní slovenských a belgických hodnôt inflácie má najnižšiu hodnotu RMSE model SETAR.

Tabuľka č. 14: Výsledné hodnoty RMSE modelov časových radov inflácie vo vybraných krajinách Európy.

Model\Krajina	Veľká					
	Slovensko	Maďarsko	Británia	Francúzsko	Belgicko	Fínsko
ARMA\ARIMA	0,273	0,457	0,261	0,279	0,443	0,078
ARFIMA	0,276	0,461	0,263	0,281	0,518	0,084
GARCH	0,281	0,474	0,244	0,286	0,466	0,077
TGARCH	0,287	0,474	0,244	0,285	0,465	0,081
GJRGARCH	0,281	0,474	0,244	0,286	0,470	0,079
AVGARCH	0,279	0,473	0,244	0,285	0,466	0,077
NAAR	0,276	0,451	0,185	0,279	0,432	0,089
SETAR	0,271	0,452	0,256	0,278	0,409	0,297
STAR	0,294	0,488	0,247	0,277	0,490	0,077
LSTAR	0,300	0,619	0,265	0,278	0,498	0,108
MSW-AR	0,318	0,319	0,125	0,166	0,482	0,081

10.4 Výsledky modelovania zamestnanosti

Charakteristika časových radov zachytávajúcích zamestnanosť sú popísané v podkapitole 8.4 tejto práce. Výsledky modelovania sme zapísali do tabuliek č. 15 a 16. Tabuľka s číslom 15 obsahuje hodnoty výberovej štandardnej odchýlky rezíduí σ_ε použitých modelov. Aj pri údajoch o zamestnanosti model MSW-AR má najnižšiu hodnotu σ_ε v každej krajine. Analyzované časové rady vývoja zamestnanosti v Maďarsku a Francúzsku potvrdili homoskedasticitu a z toho dôvodu sme nepoužili modely ARCH, GARCH a ani jeho rozšírenia. Test linearity časového radu Maďarska pre model STAR pre rôzne možnosti parametrov nezamietol nulovú hypotézu linearity modelu. Aj pri analyzovaní historických údajov zamestnanosti najnižšiu hodnotu σ_ε majú modely MSW-AR. Relatívne nízku hodnotu σ_ε majú aj jednotlivé prahové viacrežimové modely. Môžeme taktiež vidieť, že modely triedy GARCH majú najvyššie hodnoty σ_ε .

Hodnotenie predikčných vlastností ukazovateľom chyby predikcie RMSE časových radov zamestnanosti vo vybraných krajinách nie je jednoznačné. Predikcie v krajinách Slovensko, Maďarsko a Veľká Británia majú najnižšiu hodnotu RMSE použitím modelu MSW-AR. Naproti tomu v Belgicku a Fínsku model MSW-AR vykazuje najvyššie RMSE

a najnižšie vykazuje model NAAR. Môžeme pozorovať jav, že model s nízkou štandardnou odchýlkou rezíduí má pomerne vysokú chybu predikcie RMSE a naopak.

Tabuľka č. 15: Výsledné hodnoty štandardnej odchýlky rezíduí modelov časových radov vývoja zamestnanosti vo vybraných krajinách Európy.

Model\Krajina	Veľká					
	Slovensko	Maďarsko	Británia	Francúzsko	Belgicko	Fínsko
ARMA\ARIMA	14,52	16,68	86,50	48,55	8,06	9,90
ARFIMA	14,44	16,77	84,66	47,14	8,01	10,28
GARCH	14,77	NA	98,24	NA	8,16	11,71
TGARCH	14,70	NA	93,88	NA	8,11	11,70
GJRGARCH	14,59	NA	81,64	NA	8,25	11,23
AVGARCH	14,61	NA	83,48	NA	8,48	11,14
NAAR	7,51	17,27	81,33	47,24	7,68	9,73
SETAR	11,19	11,84	63,24	33,19	6,27	8,91
STAR	11,22	NA	66,65	32,85	6,07	10,19
LSTAR	12,22	16,67	68,05	42,21	7,14	10,03
MSW-AR	3,76	5,33	21,80	23,67	4,66	5,29

Tabuľka č. 16: Výsledné hodnoty chyby RMSE modelov časových radov vývoja zamestnanosti vo vybraných krajinách Európy.

Model\Krajina	Veľká					
	Slovensko	Maďarsko	Británia	Francúzsko	Belgicko	Fínsko
ARMA\ARIMA	17,63	37,61	155,21	19,96	19,59	7,98
ARFIMA	17,79	33,96	152,68	18,35	19,25	4,44
GARCH	18,11	NA	160,28	NA	19,68	4,63
TGARCH	18,03	NA	160,88	NA	19,67	4,63
GJRGARCH	17,67	NA	165,07	NA	19,36	4,61
AVGARCH	17,53	NA	168,40	NA	19,01	4,50
NAAR	31,61	36,77	148,91	29,59	18,27	4,39
SETAR	22,39	31,00	312,67	99,75	22,13	8,40
STAR	15,40	NA	157,99	118,56	19,34	4,36
LSTAR	16,65	37,95	153,44	16,05	19,86	5,25
MSW-AR	0,67	31,14	20,60	31,45	28,44	17,31

10.5 Výsledky modelovania úrokovej miery

Modelovali sme taktiež časové rady vývoja úrokovej miery v šiestich vybraných krajinách. Hodnotenie popisných vlastností na základe hodnôt výberovej štandardnej odchýlky rezíduí σ_ε je podobné ako pri ostatných dátach. V tabuľke č. 17 vidíme, že

najnižšiu σ_ε majú Markov-Switching autoregresné modely. Jediné pri modelovaní časového radu úrokovej miery v Poľsku najnižšiu hodnotu σ_ε má model GJR-GARCH. Je to nezvyčajné oproti našim skúsenostiam s inými dátami. Pri analyzovaní úrokovej miery v krajinách Švédsko, Lotyšsko a USA je už tradične spomedzi všetkých použitých modelov najväčšia hodnota σ_ε u modelov triedy GARCH a relatívne nízka je pri použití prahových modelov.

Tabuľka č. 17: Výsledné hodnoty štandardnej odchýlky rezíduí modelov časových radov úrokovej miery vybraných krajín.

Model\Krajina	Česko	Maďarsko	Poľsko	Švédsko	Lotyšsko	USA
ARMA\ARIMA	3,630	0,774	0,692	1,644	1,825	0,507
ARFIMA	3,766	0,826	0,697	1,644	1,922	0,587
GARCH	NA	0,732	0,721	1,753	2,032	0,618
TGARCH	NA	0,733	0,723	1,777	2,019	0,612
GJRGARCH	NA	0,733	0,072	1,737	2,019	0,612
AVGARCH	NA	0,729	0,726	1,751	2,046	0,624
NAAR	1,609	0,708	0,725	1,609	1,262	0,445
SETAR	1,588	0,633	0,659	1,588	1,389	0,498
STAR	1,565	0,628	0,664	0,767	1,542	0,438
LSTAR	1,473	0,654	0,698	1,508	1,615	0,503
MSW-AR	0,623	0,552	0,296	0,273	1,422	0,427

Tabuľka č. 18: Výsledné hodnoty chyby RMSE modelov časových radov úrokovej miery vybraných krajín.

Model\Krajina	Česko	Maďarsko	Poľsko	Švédsko	Lotyšsko	USA
ARMA\ARIMA	0,0409	0,1668	0,0797	0,0865	0,2780	0,1004
ARFIMA	0,0437	0,1991	0,0869	0,0927	0,2840	0,0410
GARCH	NA	0,1480	0,0871	0,0004	0,0453	0,0370
TGARCH	NA	0,1479	0,0872	0,0004	0,0454	0,0374
GJRGARCH	NA	0,1479	0,0871	0,0004	0,0452	0,0374
AVGARCH	NA	0,1490	0,0924	0,0005	0,0460	0,0376
NAAR	0,0425	0,1680	0,0985	0,0199	0,5580	0,0501
SETAR	0,0415	5,5130	0,2252	0,0062	0,1520	0,0462
STAR	0,1640	0,1396	0,1080	0,0725	0,1680	0,0753
LSTAR	0,0412	0,1466	0,0850	0,0089	0,1290	0,0467
MSW-AR	0,0548	0,0756	0,0842	0,0402	0,0459	0,0039

Chyby predikcie RMSE modelov aplikovaných na časové rady vývoja úrokovej miery v šiestich krajinách sú zobrazené v tabuľke č. 18. Teraz však modely MSW-AR nedávajú pre všetky krajiny najnižšie hodnoty, iba v prípade Maďarska a USA.

Analyzované dáta Českej republiky a Poľska majú najnižšiu chybu RMSE pri modelovaní ARMA modelom. Zaujímavé je, že v prípade modelovania časového radu úrokovej miery vo Švédsku a v Lotyšsku najnižšia chyba RMSE je pri použití modelov triedy GARCH. Keď sa však pozrieme na prislúchajúce výberové štandardné odchýlky rezíduí, vychádzajú vyššie ako pri použití zvyšných modelov. Pri hodnotení predikčných a popisných vlastností modelov aplikovaných na časové rady úrokovej miery nenachádzame zhodu.

10.6 Výsledky modelovania vývoju kurzov

Analyzovali sme aj dáta vývoja rôznych menových kurzov. Popis a charakteristika týchto dát sa nachádza v kapitole 8 tejto diplomovej práce. Hodnoty štandardnej odchýlky rezíduí σ_ε modelov použitých pri analýze spomínaných dát sú vypísané v tabuľke č. 19.

Tabuľka č. 19: Výsledné hodnoty štandardnej odchýlky rezíduí modelov časových radov vývoja vybraných menových kurzov.

Model\Mena	Koruna	Frank	Líra	Marka	Fínska marka	Šiling
ARMA\ARIMA	0,430	0,039	19,470	0,134	0,073	0,102
ARFIMA	0,450	0,042	20,560	0,143	0,075	0,108
GARCH	0,480	0,045	22,370	0,016	0,082	0,115
TGARCH	0,481	0,045	22,470	0,015	0,082	0,115
GJRGARCH	0,478	0,045	22,480	0,015	0,082	0,115
AVGARCH	0,478	0,045	22,460	0,016	0,082	0,114
NAAR	0,397	0,042	18,500	0,014	0,072	0,106
SETAR	0,426	0,038	18,800	0,013	0,069	0,092
STAR	0,396	0,036	14,640	0,012	0,065	0,094
LSTAR	0,414	0,039	18,910	0,012	0,070	0,099
MSW - AR	0,178	0,019	15,870	0,006	0,032	0,059

Z tabuľky č. 19 je jasne vidieť, že najnižšie hodnoty σ_ε majú modely MSW-AR. Najvyššiu σ_ε majú modely triedy GARCH pri všetkých menových kurzoch okrem časového radu vývoja kurzu Marka\Euro. Viacrežimové prahové modely majú pomerne nízku hodnotu σ_ε v každom prípade analyzovaného vývoja menového kurzu.

Zaujímavé na hodnotení predikčných vlastností použitých modelov na časové rady vybraných menových kurzov je, že v prípade vývoja kurzu Slovenská koruna\Euro najvyššiu hodnotu chyby predikcie RMSE má práve MSW-AR model a najnižšie RMSE majú modely triedy GARCH. Pri kurze Fínska marka\Euro najnižšia hodnota RMSE je pri

vytváraní predikcií pomocou modelu ARMA a vyššie chyby sú pri použití všetkých viacrežimových modelov. U zvyšných predpovedaných údajov menových kurzov je už najnižšia chyba RMSE pri použití modelu MSW-AR. Tie majú aj dobré popisné vlastnosti.

**Tabuľka č. 20: Výsledné hodnoty chyby RMSE
modelov časových radov úrokovej miery vybraných krajín.**

Model\Mena	Koruna	Frank	Lira	Marka	Fínska marka	Šiling
ARMA\ARIMA	0,968	0,021	8,260	0,005	0,018	0,034
ARFIMA	0,972	0,019	8,477	0,005	0,019	0,037
GARCH	0,939	0,018	8,257	0,005	0,018	0,036
TGARCH	0,939	0,018	8,256	0,005	0,018	0,036
GJRGARCH	0,946	0,018	8,246	0,005	0,018	0,037
AVGARCH	0,950	0,017	8,287	0,005	0,019	0,036
NAAR	0,943	0,018	8,524	0,006	0,019	0,043
SETAR	0,972	0,019	8,788	0,006	0,027	0,077
STAR	1,060	0,021	8,459	0,005	0,030	0,038
LSTAR	0,977	0,019	8,029	0,006	0,023	0,035
MSW - AR	1,176	0,002	1,704	0,002	0,023	0,021

10.7 Hodnotenie modelovania indikátora CISS

Posledné dáta, ktoré sme analyzovali sú časové rady hodnôt kompozitného indikátora systemického stresu CISS na rôznych finančných trhoch. Bližší popis týchto dát je v podkapitole 8.7 tejto práce. Do tabuľky č. 21 sme zapísali výsledné štandardné chyby rezíduí σ_ε modelov aplikovaných na spomínané dáta. Pri modelovaní časových radov troch trhov nám test lineariry zamietol použitie nelineárneho STAR modelu. Môžeme vidieť, že najnižšie výberové štandardné odchýlky rezíduí σ_ε vznikajú pri modeloch MSW-AR všetkých časových radov vývoja CISS. Najvyššie hodnoty σ_ε vznikajú opäť pri použití triedy GARCH modelov a ARFIMA modelov, napriek tomu, že ide o vývoj časového radu volatility, kde teória hovorí, že najlepšie by mali byť práve modely GARCH a jeho rozšírenia. Pri porovnávaní predikčných vlastností vidíme, že pri predpovedaní štyroch časových radov dostávame nízke hodnoty chyby RMSE práve pri modeloch z triedy GARCH. Iba v prípade zahraničného výmenného trhu a finančných sprostredkovateľoch najnižšie hodnoty RSME sú znova pri použití MSW-AR modelov.

Tabuľka č. 21: Výsledné štandardné odchýlky rezíduí modelov časových radov indikátora CISS rôznych finančných trhov

Model\Trh	Trh s dlhopismi	Akciový trh	Peňažný trh	Zahraničný výmenný trh	Finanční sprostredkovatelia
ARMA\ARIMA	0,0088	0,0168	0,0109	0,0161	0,0169
ARFIMA	0,0089	0,0171	0,0111	0,0162	0,0171
GARCH	0,0089	0,0172	0,0111	0,0161	0,0169
TGARCH	0,0089	0,0172	0,0112	0,0161	0,0169
GJRGARCH	0,0089	0,0173	0,0111	0,0161	0,0169
AVGARCH	0,0089	0,0173	0,0112	0,0161	0,0169
NAAR	0,0087	0,0170	0,0111	0,0160	0,0166
SETAR	0,0087	0,0165	0,0108	0,0158	0,0166
STAR	NA	NA	NA	0,0158	0,0164
LSTAR	0,0087	0,0166	0,0111	0,0159	0,0167
MSW - AR	0,0037	0,0074	0,0052	0,0067	0,0076

Tabuľka č. 22: Výsledné hodnoty chyby RMSE modelov časových radov indikátora CISS rôznych finančných trhov.

Model\Trh	Trh s dlhopismi	Akciový trh	Peňažný trh	Zahraničný výmenný trh	Finanční sprostredkovatelia
ARMA\ARIMA	0,0088	0,0159	0,0102	0,0166	0,0203
ARFIMA	0,0086	0,0156	0,0101	0,0167	0,0205
GARCH	0,0088	0,0153	0,0099	0,0165	0,0208
TGARCH	0,0088	0,0153	0,0099	0,0165	0,0208
GJRGARCH	0,0088	0,0153	0,0099	0,0165	0,0208
AVGARCH	0,0087	0,0153	0,0099	0,0166	0,0208
NAAR	0,0086	0,0162	0,0100	0,0166	0,0207
SETAR	0,0084	0,0154	0,0105	0,0167	0,0206
STAR	NA	NA	NA	0,0167	0,0206
LSTAR	0,0085	0,0157	0,0100	0,0166	0,0206
MSW - AR	0,0123	0,0263	0,0119	0,0139	0,0173

10.8 Hodnotenie výsledkov modelovania celkovo

Vo všeobecnosti z výsledkov môžeme pozorovať, že modely, ktoré mali dobré popisné vlastnosti, nie vždy dávali dobré predikcie a naopak. Po analýze výsledkov modelovania jednotlivých skupín môžeme však povedať, že najčastejšie najnižšie výberové štandardné odchýlky rezíduí a predpovednej chyby RMSE vznikali pri Markov-switching autoregresných modeloch. Rozdiely medzi výsledkami modelov MSW-AR a zvyšných modelov boli viackrát markantné. Podobne najvyššie výberové štandardné

odchýlky rezíduí a predikčné chyby RMSE mali najčastejšie modely GARCH a jeho rozšírenia. Ak by sme hodnotili len modely vytvorené v programe R, čiže by sme nebrali do úvahy výsledky MSW-AR modelov, veľmi dobré hodnoty dostávame aj pri NAAR modeloch a prahových viacrežimových autoregresných modeloch SETAR, STAR a LSTAR. Modely ARMA popřípade ARIMA len zriedkakedy mali lepšie ako priemerné výsledky.

Hypotéza o vhodnosti určitej triedy modelov pre jednu skupinu dát a vhodnosti inej triedy modelov pre inú skupinu sa na analyzovaných dátach nepotvrdila. Výsledky v jednotlivých skupinách dát sa natoľko od seba nelíšili, aby sme mohli túto hypotézu potvrdiť. Hypotéza o najlepšej triede modelov pre všetky skupiny sa viac - menej potvrdila. Modely z triedy MSW-AR majú vo väčšine prípadov najlepšie výsledky z popisného hľadiska, avšak nájdu sa časové rady, pre ktoré by bolo lepšie použiť model z inej triedy. Naším odporúčaním preto je venovať sa osobitne a dôkladne analýze každého časového radu, použiť viaceré triedy modelov na porovnanie výsledkov, a tak vybrať vhodný model. Pri ekonomických a finančných dátach nezabudnúť však na nelineárne MSW-AR modely.

ZÁVER

Počas štúdia na Fakulte matematiky, fyziky a informatiky sme mali viacero predmetov zaoberajúcich sa analýzou časových radov. Už vtedy ma táto tematika zaujala a to ma podnietilo k výberu diplomovej práce s názvom „Modelovanie ekonomických a finančných časových radov“. Navyše ma ovplyvnil aj fakt, že využitie nadobudnutých poznatkov z časových radov je rozsiahle a tematická oblasť má perspektívu i do budúcnosti, nakoľko sa v ekonómii a financiách veľmi často využíva.

Diplomová práca je rozdelená na teoretickú a aplikačnú časť. Obsahom teoretickej časti bolo zhrnutie teoretických poznatkov z analýzy časových radov potrebných na modelovanie dát a vytvorenie prehľadu lineárnych a nelineárnych modelov, ktorý by bol komplexný, štruktúrovaný a zrozumiteľný pre čitateľa s pokročilými znalosťami z matematiky. Popísaných je vyše desať modelov a ich vlastnosti.

Aplikačná časť práce zahŕňa modelovanie časových radov prezentovanými modelmi, porovnávanie ich popisných a predikčných vlastností a overenie hypotéz vyslovených v úvode diplomovej práce. Bola vybraná dostatočná vzorka dát na prijatie alebo zamietnutie stanovených hypotéz. Časových radov je 41 a sú rozdelené do 7 skupín podľa charakteru dát. Porovnávanie popisných a predikčných vlastností modelov nie je jednoduché, nakoľko počet modelov aplikovaných na jednotlivé časové rady je desať. Na porovnanie popisných vlastností modelu sme vybrali hodnotiaci ukazovateľ výberová štandardná odchýlka rezíduí, ktorá vzniká pri analýze dát daným modelom. A na porovnanie predikčných vlastností modelov sme zvolili ako ukazovateľ predpovednú chybu RMSE predikcie, ktorú konfrontujeme s reálnou hodnotou časového radu. Vyjadrenia o platnosti stanovených hypotéz robíme na základe výsledkov modelovania jednotlivých skupín dát.

Vyjadrenia o prijatí alebo zamietnutí skúmaných hypotéz sa nachádzajú na konci diplomovej práce. Po ukončení analytickej práce sme zistili, že pre vybrané skupiny dát neexistuje trieda modelov s najlepšimi popisnými aj predikčnými vlastnosťami. Môžeme povedať, že najlepšie popisné vlastnosti má vo väčšine prípadov model z triedy MSW-AR. Model z tejto triedy mal často aj najlepšie predikčné vlastnosti. Okrem tejto triedy modelov stabilne dobré výsledky mali aj ostatné nelineárne modely. A dá sa povedať, že

lineárne modely ARMA, ARIMA a ARFIMA jemne zaostávali za ostatnými, až na zvyčajne nepriaznivo hodnotené modely triedy GARCH a jeho rozšírení.

Avšak tvrdenie, že model z triedy MSW-AR je univerzálny pre všetky druhy ekonomických a finančných časových radov by bolo príliš trúfalé. Záverom je, že každý časový rad, bez ohľadu na to akého je charakteru, si žiada dôkladnú analýzu a výber vhodného modelu závisí na viacerých faktoroch, avšak vyskúšať použitie MSW-AR modelu odporúčam. Zistením je aj to, že jeden model môže byť lepší na popis časového radu a druhý na vytváranie predpovedí.

Cieľom diplomovej práce bolo modelovať ekonomické a finančné časové rady pomocou rôznych lineárnych a nelineárnych, najmä viacrežimových, modelov a porovnať ich popisné a predikčné vlastnosti. Po ukončení modelovania a zhodnotení výsledkov môžeme povedať, že stanovený cieľ sme splnili a zistenia práce môžu byť prínosom. Pre mňa sú získané poznatky nielen teoretické, ale hlavne praktické. Verím, že ich budem môcť využiť aj v budúcnosti, pretože analyzovanie časových radov, práca v programe R a samotná práca s dátami bola pre mňa nielen povinnosťou, ale aj záujmom. Prostredie programu R s využitím R Studia je vhodné nielen pre informatikov a jeho používanie na analýzu časových radov môžem iba odporučiť. Späťne môžem s určitosťou povedať, že svoje rozhodnutie vybrať si diplomovú prácu s danou témou neľutujem a chcela by som nadobudnuté poznatky ďalej rozvíjať.

V ďalšom výskume by bolo možné zamerať sa konkrétnejšie na určitý typ dát, napríklad vývoj jedného makroekonomického ukazovateľa vo väčšom počte krajín a porovnávať vlastnosti modelov pomocou viacerých kritérií. Je otázkou, či by výsledky boli iné pri výbere odlišných kritérií hodnotenia modelov. Ďalší výskum by sa mohol zamerať hlavne na predikciu a prísnu minimalizáciu viacerých typov predpovedných chýb. Inou alternatívou je modelovanie časových radov v prostredí E-views a porovnanie nadobudnutých výsledkov pomocou dvoch rôznych programov. Takéto rozšírenie svojej práce by som si vybrala ja, ale ako vidno, možností pri rozvíjaní tejto témy je viacero.

ZOZNAM BIBLIOGRAFICKÝCH ODKAZOV

- [1] ANDERSON, O. *Time series analysis and forecasting: the Box-Jenkins approach*. London: Butterworth, 1976. 182 p. ISBN 0408706759.
- [2] ARLT, J. *Moderní metody modelování ekonomických časových řad*. Praha: Grada Publishing, 1999. 307 s. ISBN 80-7169-539-4.
- [3] BHARDWAJ, G., SWANSON, N. R. An Empirical Investigation of the Usefulness of ARFIMA Models for Predicting Macroeconomic and Financial Time Series. In *Journal of Econometrics*. Amsterdam: Elsevier, 2006. ISSN 0304-4076. Vol. 131, No. 1-2, pp. 539-578.
- [4] BOLLERSLEV, T. Generalized Autoregressive Conditional Heteroskedasticity. In *Journal of Econometrics*. Amsterdam: Elsevier, 1986. Vol. 31. pp. 307-327.
- [5] BOX, G.E.P., JENKINS, G. *Time Series Analysis, Forecasting and Control*. San Francisco: Holden Day, 1970. 575 p. ISBN 0-470-27284-8.
- [6] BOX, G. E. P., PIERCE, D. A. Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models. In *Journal of the American Statistical Association*. Alexandria: American Statistical Association, 1970. Vol. 65. No.332. pp. 1509–1526.
- [7] BROCK, W.A., DECHERT, W. D., SCHEINKMAN, J. A. *A Test for Independence Based on the Correlation Dimension*. Madison: Social Systems Research Institute, 1987. ISBN:0-521-77297-4. p.18.
- [8] BROCKWELL. P.J., DAVIS, R.A. *Time Series: Theory and Methods*. Berlin: Springer-Verlag, 1986. 584 p. ISBN 0-387-96406-1.
- [9] CIPRA, T. *Analýza časových řad s aplikacemi v ekonomii*. Praha: SNTL – Nakladatelství technické literatury, 1986. 246 s.

- [10] DIEBOLD, F.X., KILIAN, L., NERLOVE, M. Time Series Analysis. In BLUME, L., DURLAUF, S. *The New Palgrave Dictionary of Economics, Second Edition*. London: MacMillan, 2006. 2650 pp. ISBN 9781137336583.
- [11] DING, Z., GRANGER, C.V.J., ENGLE, R.F. A long memory property of stock market returns and a new model. In *Journal of Empirical Finance*. Amsterdam: Elsevier, 1993. ISSN 0927-5398. Vol. 1, pp. 83-106.
- [12] ENGLE, R. F. Autoregressive Conditional Heteroskedasticity With Estimates of the Variance of U.K. Inflation. In *Econometrica*. Yale: The Econometric Society, 1982. ISSN 00129682. Vol. 50, No. 4, pp. 987-1007.
- [13] FISCHER, B. *Decompositions of Time Series. Comparing Different Methods in Theory and Practice*. Luxembourg: Eurostat, 1995. 68 p.
- [14] FOUSKITAKIS, G. N., FASSOIS, S. D. *On the estimation of long-memory time series models*. Patras: University of Patras in Greece, 2000.
- [15] FRANSES, P.H., DIJK, D. *Nonlinear Time Series Models in Empirical Finance*. Cambridge: Cambridge University Press, 2003. 280 p. ISBN 0-511-01100-8.
- [16] GLOSTEN, L. R. J., JAGANNATHAN, R., RUNKLE, D. E.,. On the Relation between the Expected Value and the Volatility of Nominal Excess Return on stocks. In *Journal of Finance*. Amsterdam: Elsevier, 1993. ISSN 0927-5398. Vol. 48. No. 5. pp. 1779-1801.
- [17] HAMILTON, J. D. A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. In *Econometrica*. Yale: The Econometric Society, 1989. Vol. 57, pp. 357-384.
- [18] HOLLÓ, D., KREMER, M., DUCA, M. CISS – A Composite Indicator of Systemic Stress in the Financial System. In *Working Paper Series*. Frankfurt am Main: European Central Bank, 2012. ISSN 1725-2806. No. 1426. p. 51.

- [19] CHATFIELD, C. *The Analysis of Time Series: An Introduction. Sixth Edition.* Taylor&Francis, 2003. 352 p. ISBN 1584883170.
- [20] JARQUE, C.M., BERA, A. K. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. In *Economics Letters*. Canberra: Australian National University, 1980. ISSN 0165-1765. Vol. 6. No. 3. pp. 255–259.
- [21] KIRCHGÄSSNER, G., Wolters, J. *Introduction to Modern Time Series Analysis.* Berlin: Springer, 2007. 274 p. ISBN 978-3-540-68735-1.
- [22] KLEIN, J. L. *Statistical Visions in Time: A History of Time Series Analysis, 1662-1938.* Cambridge: Cambridge University Press, 2005. 345 p. ISBN 0521023173.
- [23] KOZAK, J., ARLT, J., HINDLS, R. *Úvod do analýzy ekonomických časových řad.* Praha: Vysoká škola ekonomická, 1994. 208 s. ISBN 80-7079-760-6.
- [24] LEBARON, B. Forecast improvements using a volatility index. In *Journal of Applied Econometrics*. New Jersey: Wiley, 1992.vol. 7. No. S1. pp 137-149.
- [25] LJUNG, G.M., BOX, G.E.P. On a Measure of a Lack of Fit in Time Series Models. In *Biometrika*. London: Biometrika Trust, 1978. ISSN 1464-3510. Vol. 65. No. 2. pp 297–303.
- [26] NERLOVE, M., GREYER, D. M., CARVALHO, J. L. *Analysis of Economic Time Series: A Synthesis.* New York: Academic Press, 1979. 468 p. ISBN0125157509.
- [27] PRIESTLY, M.B. *Nonlinear and Non-Stationary Time Series.* New York: Academic Press, 1988. 227 p. ISBN 012564910X.
- [28] SAID, S. E., DICKEY, D.A. Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order. In *Biometrika*. London: Biometrika Trust, 1984. ISSN 1464-3510. Vol. 71. No. 3. pp. 599–607.

- [29] SCHWERT, G.W. Stock volatility and the crash of '87. In *Review of Financial Studies*. Oxford: Oxford University Press, 1990. Vol. 3. No. 1. pp. 77-102.
- [30] SWANSON, N. R., BHARDWAJ, G. An Empirical Investigation of the Usefulness of ARFIMA Models for Predicting Macroeconomic and Financial Time Series. In *Journal of Econometrics*. Amsterdam: Elsevier, 2006. Vol. 131. No. 1-2. pp. 539-578.
- [31] TAYLOR, S.J. *Modelling financial time series*. Chichester: Wiley, 1986. p. 268. ISBN 9812770844.
- [32] TERÄSVIRTA, T. *Handbook of Financial Time Series*. Springer, 2009. 1050 p. ISBN 978-3-540-71296-1.
- [33] TONG, H. On a threshold model. In Chen, C, (ed.) *Pattern recognition and signal processing*. NATO ASI Series E: Applied Sc. Amsterdam: Sijthoff and Noordhoff, 1978. ISBN 9789028609785. No. 29. pp. 575-586.
- [34] TSAI, H., CHAN, K. A note on inequality constraints in the GARCH model. In *Econometric Theory*. Cambridge: Cambridge University Press, 2008. Vol. 24. No. 03. pp 823-828.
- [35] TSAY, R. *Analysis of financial time series*. New Jersey: Wiley., 2005. 638 p. ISBN 10 0-471-69074.
- [36] TSAY, R., CHEN, R. Nonlinear additive ARX models, *Journal of the American Statistical Association*. Texas: Taylor&Francis, 1993. Vol. 88, No. 423, pp. 955-967.
- [37] VERBEEK, M. *A guide to modern econometrics*. New Jearsey: Wiley, 2008. 472 p. ISBN 0-470-51769-7.

- [38] WHITE, H. A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. In *Econometrica*. Yale: The Econometric Society, 1980. ISSN 00129682. Vol. 48, No. 4, pp. 817-838.
- [39] WOLD, H. *A study in the analysis of stationary time series*. Stockholm: Almqvist&Wiksell, 1938. 236 p.
- [40] WU, J. *Threshold GARCH Model: Theory and Application*. Ontario: The University of Western Ontario, 2010. 42 p.
- [41] Yang, Y. *Can the strengths of AIC and BIC be shared? A conflict between model identification and regression estimation*. In *Biometrika*. London: Biometrika Trust, 2005. ISSN 1464-3510. Vol. 92. No. 4. pp. 937-950.
- [42] ZAKOIAN, J.M. Threshold heteroskedastic models. In *Journal of Economic Dynamics and Control*. Amsterdam: Elsevier, 1994. ISSN: 0165-1889. Vol. 18. No. 5. pp. 931-955.
- [43] R Development Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing: Vienna, 2012. ISBN 3-900051-07-0. Dostupné na internete: <<http://www.R-project.org/>>.
- [44] Európska komisia. Meranie inflácie v EÚ. [cit. 15.04.2013] Dostupné na internete: <http://ec.europa.eu/economy_finance/focuson/inflation/measuring_sk.htm>
- [45] European Commission. Statistics Database. [cit. 10.03.2013] Dostupné na internete: <http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database>
- [46] Štatistický úrad Slovenskej republiky. Metodické vysvetlivky. [cit. 15.04.2013] Dostupné na internete: <<http://portal.statistics.sk/showdoc.do?docid=1648>>
- [47] European Central Bank. Statistical data warehouse. [cit. 18.03.2013] Dostupné na internete: <<http://sdw.ecb.europa.eu/>>

PRÍLOHA 1

Táto príloha obsahuje príkazy použité na modelovanie časových radov v programe R aj s popisom.

```
Inflacia_Finsko<-read.table(file="D:\\DIPL ZUZKA\\Data\\HICP
Finsko.tIFint", header=T) # načítanie dát s hlavičkou
„I_Finsko“
attach(Inflacia_Finsko) # pripojenie dát
IFinsko<-ts(I_Finsko, frequency=12,start=c(1996,1))#
definovanie časového radu s mesačnou periódou a začiatkom
v januári 1996
plot.ts(IFinsko) # vykreslenie grafu časového radu
HICP_dekompozicia <- stl(IFinsko, t.window=10, s.window=4,
robust=TRUE) # dekompozícia časového radu
plot(IFinsko_dekompozicia) # vykreslenie zložiek časového
radu
IFin<-IFinsko_dekompozicia$time.series[,3] # rezíduá po
dekompozícii
n <- length(IFin) # dĺžka časového radu
adf.test(IFin) # test stacionarity: H0 - data majú jednotkový
koreň
#IFinDif<-diff(IFin) # v prípade nestacionarity,
diferencovanie časového radu
#adf.test(IFinDif) # opätovné overenie stacionarity
acfGDP<-acf(IinN,100) # vykreslenie ACF
pacfGDP<-pacf(IFin,100) # vykreslenie PACF
NaPredpoved<-ts(IFin[1:(n-13)], frequency=12,start=c(1996,1))
# vytvorenie časového radu bez posledných hodnôt na
predpovedanie
Real <- ts(IFin[(n-12):(n)],frequency=12,start=c(2012,1)) #
vytvorenie časového radu reálnych hodnôt, ktoré budeme
porovnávať s predpoveďami

# ARMA alebo ARIMA model
IFin.arma <- arima(IFin, order = c(1,0,0)) # ARMA(1,0) model
summary(IFin.arma) # vlastnosti modelu, vrátane AIC a BIC
kritérií
Box.test(IFin.arma$residuals, lag = 1, type = c("Box-Pierce",
"Ljung-Box"), fitdf = 0) # test nezávislosti rezíduí: H0 -
rezíduá sú nezávislé
bds.test(IFin.arma$residuals, m=2) # BDS test nezávislosti
a identickosti rozdelenia rezíduí modelu: H0 - rezíduá sú
nezávislé a i.i.d.
jarque.bera.test(IFin.arma$residuals) # test normality
rezíduí: H0 - rezíduá sú normálne rozdelené
sd(IFin.arma$residuals) # štandardná odchýlka modelu
```

```

# Vytvorenie a porovnávanie predpovedí
NaPredpoved.arma <- arima(NaPredpoved, order = c(1,0,0)) #
modelovanie časového radu na predpoveď
predpoved<-forecast(NaPredpoved.arma , h=13) # vytvorenie
predpovede 13-tich hodnôt a intervalu spoľahlivosti
plot(predpoved) # vytvorenie grafu predpovede
lines(Real) # dokreslenie reálnych hodnôt
plot(predpoved,IFinlim=c(2012.1,2013)) # priblíženie grafu na
predpovede
lines(Real) # dokreslenie reálnych hodnôt
mod<-predpoved$mean # hodnoty predpovede
sqrt(sum((Real-mod)^2)/14) # výpočet odmocniny priemernej
štvorcovej chyby

# ARFIMA model
IFin.arfima <- arfima(IFin, order=c(0,0,1)) # ARFIMA(0,0,1)
model
summary(IFin.arfima) # vlastnosti modelu vrátane AIC a BIC
kritérií
rez<-resid(IFin.arfima) # definovanie rezíduí ARFIMA modelu
bds.test(rez[[1]],m=2) # BDS test rezíduí modelu
Box.test(rez[[1]], lag = 1, type = c("Box-Pierce", "Ljung-
Box"), fitdf = 0) # test nezávislosti rezíduí
jarque.bera.test(rez[[1]]) # test normality rezíduí
sd(rez[[1]]) # štandardná odchýlka modelu

# Vytvorenie a porovnávanie predpovedí
IFin.arfima <- arfima(NaPredpoved) # modelovanie časového
radu na predpoveď
predpoved<-forecast(NaPredpoved.arfima , h=13) # vytvorenie
predpovedí
plot(predpoved) # vytvorenie grafu predpovede
plot(predpoved,IFinlim=c(2012.1,2013)) # priblíženie grafu na
predpovede
lines(Real) # dokreslenie reálnych hodnôt
mod<-predpoved$mean # hodnoty predpovede
sqrt(sum((Real-mod)^2)/14) # výpočet odmocniny priemernej
štvorcovej chyby

# Ďalšie diagnostické testy
acf(IFin.arma$residuals,60) # vytvorenie ACF rezíduí
ARMA\ARIMA modelu
acf(IFin.arma$residuals^2,100) # vytvorenie druhých mocnín
ACF rezíduí ARMA\ARIMA modelu na posúdenie homoskedasticity
acf(rez[[1]],60) # vytvorenie ACF rezíduí ARFIMA modelu
acf(rez[[1]]^2,100) # vytvorenie druhých mocnín ACF rezíduí
ARFIMA modelu na posúdenie homoskedasticity

```

```

# ARCH, GARCH model
mod.garch<- garchFit(~arma(3,1)+garch(1,1), data = IFin,
cond.dist=c("sstd"), include.mean=F,trace=FALSE, delta =
2,leverage = FALSE) # vytvorenie GARCH modelu
summary(mod.garch)
bds.test(mod.garch@residuals)
sd(mod.garch@residuals)
# Vytvorenie a porovnávanie predpovedí
predp.garch<- garchFit(~arma(3,1)+garch(1,1), data =
NaPredpoved, cond.dist=c("sstd"), include.mean=F,trace=FALSE,
delta = 2,leverage = FALSE)
predpoved<-predict(predp.garch, n.ahead = 13,mse="uncond")[1]
sqrt(sum((Real-predpoved$meanForecast)^2)/14)

# TGARCH model
mod.tgarch <- garchFit(~arma(3,1)+aparch(1,1), data = IFin,
cond.dist=c("sstd"), include.mean=F,trace=FALSE, delta =
1,leverage = TRUE) # vytvorenie TGARCH modelu
summary(mod.tgarch)
bds.test(mod.tgarch@residuals)
sd(mod.tgarch@residuals)
# Vytvorenie a porovnávanie predpovedí
predp.tgarch<- garchFit(~arma(3,1)+garch(1,1), data =
NaPredpoved, cond.dist=c("sstd"), include.mean=F,trace=FALSE,
delta = 1,leverage = TRUE)
predpoved<-predict(predp.tgarch, n.ahead =
13,mse="uncond")[1]
sqrt(sum((Real-predpoved$mean)^2)/14)

# GJR-GARCH model
mod.gjrgarch<- garchFit(~arma(3,1)+aparch(1,1), data = IFin,
cond.dist=c("sstd"), include.mean=F,trace=FALSE, delta =
2,leverage = TRUE,include.delta=FALSE) # vytvorenie GJR-GARCH
modelu
summary(mod.gjrgarch)
bds.test(mod.gjrgarch@residuals)
sd(mod.gjrgarch@residuals)
# Vytvorenie a porovnávanie predpovedí
predp.gjrgarch<- garchFit(~arma(3,1)+garch(1,1), data =
NaPredpoved, cond.dist=c("sstd"), include.mean=F,trace=FALSE,
delta = 2,leverage = TRUE,include.delta=FALSE)
predpoved<-predict(predp.gjrgarch, n.ahead =
13,mse="uncond")[1]
sqrt(sum((Real-predpoved$mean)^2)/14)

# AVGARCH model
mod.avgarch<- garchFit(~arma(3,1)+aparch(1,1), data = IFin,

```



```

cond.dist=c("sstd"), include.mean=TRUE, trace=FALSE, delta =
1,leverage = TRUE,include.delta=FALSE) # vytvorenie AVGARCH
modelu
summary(mod.avgarch)
bds.test(mod.avgarch$residuals)
sd(mod.avgarch$residuals)
# Vytvorenie a porovnanie predpovedí
predp.avgarch<- garchFit(~arma(3,1)+garch(1,1), data =
NaPredpoved, cond.dist=c("sstd"), include.mean=TRUE,
trace=FALSE, delta = 1,leverage = TRUE,include.delta=FALSE)
predpoved<-predict(predp.avgarch, n.ahead = 13,mse="uncond")
sqrt(sum((Real-predpoved$meanForecast)^2)/14)

# NAAR model
mod.aar <- aar(IFin, m=6,d=6) # vytvorenie NAAR modelu
s dimenziou m=6 a časovým omeškaním d=6
summary(mod.aar)
bds.test(mod.aar$residuals[-c(1,2)],m=2)
jarque.bera.test(mod.aar$residuals)
sd(mod.aar$residuals)
# Vytvorenie a porovnanie predpovedí
predp.aar<- aar(NaPredpoved, m=6,d=6)
predpoved<-predict(predp.aar, n.ahead = 13)
sqrt(sum((Real-predpoved)^2)/14)
plot(Real)
lines(predpoved, col='blue')

# SETAR model
setarTest(IFin, m=6, thDelay=5,nboot=1,test="1vs") # 2 testy
linearity pri zvolených parametroch:
# 1. Test: H0 - 1 režim lepší ako 2 režimy
# 2. Test: H0 - 1 režim lepší ako 2 3 režimy
setarTest(IFin, m=6, thDelay=5,nboot=1,test="2vs3")# test 2
režimy vs. 3 režimy: H0 - 2 režim lepšie ako 3 režimy
mod.setar <- setar(IFin, m=6, thDelay=1,nthresh=2) #
vytvorenie SETAR modelu s dimenziou m=6, časovým omeškaním
thDelay=1 a počtom prahov=2 (čiže 3 režimy)
plot(mod.setar) # vytvorenie grafu rozdelenia hodnôt do
jednotlivých režimov
summary(mod.setar)
bds.test(mod.setar$residuals,m=2)
jarque.bera.test(mod.setar$residuals)
sd(mod.setar$residuals)
# Vytvorenie a porovnanie predpovedí
predp.setar<- setar(NaPredpoved, m=6, thDelay=5,nthresh=2)
predpoved<-predict(predp.setar, n.ahead = 13)
sqrt(sum((Real-predpoved)^2)/14)
plot(predpoved, col='blue')
lines(Real)

```

```

# STAR model
mod.star <- star(IFin, m=5, thDelay=3, noRegimes=3) #
vytvorenie STAR modelu s m=5, thDelay=3 a maximálnym počtom
režimov 3 (test linearity a test 1 vs. 2 režimy sa urobí
automaticky)
summary(mod.star)
bds.test(mod.star$residuals[-c(1,2)],m=2)
jarque.bera.test(mod.star$residuals)
sd(mod.star$residuals[-c(1,2)])
# Vytvorenie a porovnávanie predpovedí
predp.star<- star(NaPredpoved, m=5, noRegimes=3, thDelay=3)
predpoved<-predict(predp.star, n.ahead = 13)
sqrt(sum((Real-predpoved)^2)/14)
plot(Real)
lines(predpoved, col='blue')

# LSTAR model
mod.lstar <- lstar(IFin, m=6, thDelay=5,
control=list(maIFinit=1000)) # vytvorenie STAR modelu s m=5,
thDelay=3 a maximálnym počtom režimov 3 (test linearity sa
urobí automaticky)
summary(mod.lstar)
bds.test(mod.lstar$residuals[-c(1,2)],m=2)
sd(mod.lstar$residuals[-c(1,2)])
plot(mod.lstar)
# Vytvorenie a porovnávanie predpovedí
predp.lstar<- lstar(NaPredpoved,m=6, thDelay=5,
control=list(maIFinit=1000))
predpoved<-predict(predp.lstar, n.ahead = 13)
sqrt(sum((Real-predpoved)^2)/14)
plot(predpoved, col='blue', ylim=c(-0.2,0.25))
lines(Real)

```