

**ČESKÁ ZEMĚDĚLSKÁ UNIVERZITA V PRAZE**  
**PROVOZNĚ EKONOMICKÁ FAKULTA**

**Prognostické modely v oblasti modelování finančních  
časových řad**

disertační práce

Autor: RNDr. Vladimíra PETRÁŠKOVÁ  
Školitel: Doc. RNDr.Bohumil KÁBA,CSc., katedra statistiky

Praha 2006

## Obsah

<b>Úvod .....</b>	<b>5</b>
<b>1. Přehled o současném stavu problematiky .....</b>	<b>7</b>
<b>1.1. Finanční časové řady. Akciový trh .....</b>	<b>7</b>
1.1.1. Základní pojmy .....	7
1.1.2. Vlastnosti finančních časových řad .....	9
1.1.3. Akciový trh .....	11
1.1.3.1. Základní pojmy .....	11
1.1.3.2. Výpočet indexů .....	13
1.1.3.3. Přehled nejznámějších indexů .....	14
1.1.3.4. Akciové analýzy .....	18
<b>1.2. Box-Jenkinsova metodologie .....</b>	<b>21</b>
1.2.1. Box-Jenkinsova metodologie: Stacionární procesy .....	21
1.2.1.1. Základní pojmy a aparát Box-Jenkinsovy metodologie.....	21
1.2.1.2. Lineární proces .....	26
1.2.1.2.1. Proces klouzavých součtů $MA(q)$ .....	27
1.2.1.2.2. Autoregresní proces $AR(p)$ .....	28
1.2.1.2.3. Smíšený proces $ARMA(p,q)$ .....	29
1.2.1.3. Výstavba modelu $ARMA(p,q)$ .....	29
1.2.2. Boxova-Jenkinsova metodologie: Nestacionární procesy .....	32
1.2.2.1. Integrovaný smíšený model $ARIMA(p,d,q)$ .....	32
1.2.2.2. Modely časových řad s dlouhou pamětí $ARFIMA(p,d,q)$ .....	34
<b>1.3. Modely s přenosovou funkcí .....</b>	<b>38</b>
1.3.1. Základní pojmy .....	38
1.3.2. Základní charakteristika modelů s přenosovou funkcí .....	40
1.3.2.1. Přenosová funkce .....	40
1.3.2.2. Koyckův model a množina racionálních modelů .....	41
1.3.2.3. Nejčastější tvary přenosové funkce .....	45
1.3.2.4. Vícenásobný vstup .....	46
1.3.3. Identifikace modelu .....	48

1.3.3.1. Předběžné kroky .....	49
1.3.3.2. Identifikace modelu metodou LTF .....	51
1.3.3.3. Identifikace na základě vzájemného vztahu CCF a vah přenosové funkce .....	56
1.3.4. Odhady parametrů modelu .....	59
1.3.5. Ověřování modelu .....	62
1.3.5.1. Obecná metoda ověřování modelu .....	62
1.3.5.2. Obecné zásady výstavby a ověření modelu .....	66
1.3.6. Předpovědi .....	67
<b>1.4. Modely volatility .....</b>	<b>72</b>
1.4.1. Lineární modely volatility – přehled .....	73
1.4.1.1. Modely ARCH .....	74
1.4.1.2. Modely GARCH .....	76
1.4.2. Výstavba modelů volatility .....	80
<b>2. Cíl práce a metodika zpracování .....</b>	<b>84</b>
3.1. Cíl práce .....	84
3.1. Metodika zpracování .....	85
2.2.1. Určení vhodného modelu, který popisuje závislost indexu PX 50 na oborových indexech a modelu popisující závislost indexu PX 50 a indexu eurozóny .....	86
2.2.2. Určení nejvhodnějšího modelu pro míru zisku indexu PX 50 .....	87
2.2.3. Určení jednodenní předpovědi volatility míry zisku a určení předpovědi roční volatility míry zisku .....	88
<b>3. Výsledky disertační práce .....</b>	<b>90</b>
3.1. Index PX 50 versus oborové indexy .....	91
3.2. Index PX 50 versus index Dow Jones EURO STOXX 50 .....	101
3.3. Modelování indexu PX 50 pomocí modelů volatility .....	121
3.3.1. Výběrové charakteristiky a rozdělení míry zisku indexu PX 50 .....	121
3.3.2. Model volatility .....	124
3.4. Předpověď volatility pro míru zisku indexu PX 50 a indexu Standard&Poo'rs500 .....	129

<b>Závěr .....</b>	<b>138</b>
<b>Příloha .....</b>	<b>141</b>
Výstup z programu SCA .....	142
Výstup z programu MATHEMATICA .....	162
Tabulka č.1 .....	166
Tabulka č.2 .....	169
<b>Literatura.....</b>	<b>172</b>

## Úvod

Analýza časových řad se v posledních letech stala velmi dynamicky rozvíjející disciplínou. Vznikla řada nových efektivních a netradičních postupů a metod modelování časových řad. V současnosti není možné provádět důležitá ekonomická a finanční rozhodnutí bez důkladné analýzy vývoje základních ukazatelů. Velký důraz se klade nejen na konstrukci statisticko-ekonometrických modelů charakterizujících základní rysy vývoje celého hospodářství, ale i na konstrukci modelů finančních časových řad popisujících chování finančního trhu. V současné době časové řady ukazatelů hospodářského vývoje a finančního trhu začínají mít jistou informační schopnost. Správná aplikace současných metod analýzy časových řad může vést k informacím, které usnadňují rozhodovací činnost na různých úrovních národohospodářského řízení.

V 90. letech 20. století se začal formovat v České republice tržní systém a jednou z oblastí, která se dostala do popředí zájmu, je oblast finančnictví, bankovníctví a burz. Burzovníctví je vysoce specializovanou oblastí, která bezesporu hraje základní roli v každé vyspělé tržní ekonomice. Bez dobře fungujících burz by neexistoval kapitálový trh, který má za úkol shromažďovat volný kapitál a orientovat jej tam, kde ho lze co možná nejúčelněji využít. Náš kapitálový trh v porevolučních letech vznikl současně s prvními emisemi dluhopisů. Dalším obrovským počinem v tomto směru se stalo zahájení činnosti Burzy cenných papírů v Praze, kde se začalo obchodovat s akciemi akciových společností vzešlých z privatizace. Přestože od té doby došlo již k několika významným změnám a překonalo se nespočet překážek, dá se říci, že náš kapitálový trh je oproti kapitálovým trhům západních zemí stále ještě na počátku a bude ještě dlouho trvat, než se dostaneme na úroveň vyspělých zemí moderního světa. Zřejmě nejdůležitějším ukazatelem vývoje celé burzy cenných papírů byl od roku 1994 do roku 2006 index PX 50. Předkládaná práce se zabývá sledováním různých faktorů, které mají dopad na vývoj tohoto nejsledovanějšího indexu České republiky.

Při modelování indexu PX 50 se vychází ze základních metod modelování finančních časových řad. Základ je tvořen Box-Jenkinsovou metodologií výstavby jak stacionárních modelů ARMA, tak nestacionárních modelů ARIMA a ARFIMA, jež tvoří základ pro další

modelování. Z Box – Jenkinsovy metodologie vychází modely s přenosovou funkcí. Tyto modely pokrývají poměrně širokou oblast časových řad. Umožňují sledovat závislost mezi časovými řadami a tvorbu modelů, které tyto závislosti popisují. Tyto modely jsou v práci využity při sledování závislosti indexu PX 50 a oborových indexů, popř. indexu Dow Jones EURO STOXX 50. Vzhledem k tomu, že užití lineárních modelů při modelování finančních časových řad je značně problematické, a to z důvodu, že tyto modely nejsou schopny zachytit proměnlivosti podmíněného rozptylu, jsou v práci stručně zmíněny základy modelů volatility. Tyto modely se zdají být v současnosti nejvhodnějšími modely pro modelování finančních časových řad.

S novými směry rozvoje analýzy časových řad dochází také ke vzniku nového software, který zpřístupňuje tyto metody stále širšímu okruhu zájemců. Je však třeba poznamenat, že každý software řeší danou problematiku trochu jinak a není tak neobvyklé, že aplikace stejně nazvaných procedur v různých programových produktech může vést k nestejným výsledkům. V předkládané práci byly použity program SCA (při tvorbě modelů s přenosovou funkcí) a program Eviews 5 (při tvorbě modelů volatility). Další použité programy jsou STATISTICA a MATHEMATICA.

# 1. Přehled o současném stavu problematiky

Předkládaná disertační práce se zabývá problematikou modelování finančních časových řad. Tyto časové řady jsou specifické a ve srovnání s ostatními typy ekonomických časových řad mají určité rysy, které vyžadují netradiční přístupy k jejich analýze. Analýza finančních časových řad je relativně nová disciplína. Aplikace jejich metod vede k informacím, které jsou klíčové nejen pro různé finanční analýzy, ale i pro investiční společnosti či obchodníky s cennými papíry.

## 1. 1. Finanční časové řady. Akciový trh .

Vzhledem k tomu, že se tato práce zabývá modelováním burzovního indexu PX 50, který od roku 1994 do března 2006 patřil v České republice mezi nejznámější a nejsledovanější index, je nutné se blíže seznámit se základními pojmy, které se týkají obchodování na burze a charakteristiky finančních časových řad.

### 1.1.1. Základní pojmy

#### Časová řada

Časová řada je statistická časová řada, jejíž chování je zatíženo nejistotou, na rozdíl od deterministické časové řady, jejíž chování lze striktně popsat matematickým vzorcem, takže lze např. zkonstruovat její přesnou předpověď. Časovou řadu lze chápat jako konkrétní realizaci náhodného procesu.

#### Značení

Pozorování v časové řadě budeme značit velkými písmeny z konce abecedy ( $X, Y, Z$ ) a indexovat malým písmenem  $t$ , které bude představovat čas. Časovou řadu budeme zapisovat ve tvaru  $X_1, X_2, \dots, X_n$  nebo analogicky  $X_t$ , kde  $t = 1, 2, \dots, n$ .

#### Finanční trh

Finanční trh je součástí tržního systému, který představuje nabídku a poptávku peněz a kapitálu. Existují tři druhy finančních trhů, a to dluhopisové trhy, akciové trhy a devizové trhy. Na dluhopisových trzích se obchoduje s dluhovými cennými papíry, přičemž základní informací tohoto trhu je cena dluhopisu. Na akciovém trhu se obchoduje

s akciemi a základní informací je cena akcie. Na devizovém trhu se obchoduje s peněžními prostředky v různých měnách a směrodatná je cena měny.

### **Finanční časové řady**

Finanční časové řady jsou řady, které podávají kvantitativní informace o finančním trhu. Ceny produktů finančního trhu jsou sledovány v určité časové frekvenci a tvoří tak časovou řadu. Řady, které vycházejí z cen nebo které charakterizují ceny a jejich vývoj, se označují jako finanční časové řady. Cenové změny mohou být definovány různým způsobem, přičemž často se používá relativní cenová změna jednoduše označovaná jako míra zisku.

### **Diskrétní míra zisku**

Diskrétní míra zisku v čase  $t$  (značíme  $R_t$ ) je relativní změna ceny z času  $t-1$  do času  $t$  (obvykle udávaná v procentech)

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}}, \quad (1.1.1)$$

kde  $P_t$  je náhodná veličina představující cenu příslušného finančního aktiva v čase  $t$ , který je měřen ve zvolených časových jednotkách (dnech, týdnech, měsících apod.).

### **Logaritmická míra zisku**

Logaritmická míra zisku (logaritmická cenová změna) v čase  $t$  (značíme  $r_t$ ) je přirozený logaritmus úročitele  $1 + R_t$  odpovídajícího diskrétní míře zisku  $R_t$

$$r_t = \ln(1 + R_t) = \ln\left(\frac{P_t}{P_{t-1}}\right) = p_t - p_{t-1}, \quad (1.1.2)$$

kde symbol  $p_t$  označuje logaritmus ceny

$$p_t = \ln(P_t). \quad (1.1.3)$$



V praxi diskrétní a logaritmická míra zisku většinou nabývají podobných numerických hodnot, neboť pro hodnoty  $R_t$  nepříliš vzdálené od nuly lze pomocí Taylorova rozvoje ([57]) aproximovat

$$r_t = \ln(1 + R_t) \approx R_t. \quad (1.1.4)$$

Z tohoto důvodu budeme v této práci jednotně používat symbol  $r_t$  a mluvit jednoduše o míře zisku.

### 1.1. 2. Vlastnosti finančních časových řad

Základní odlišnost finanční časové řady od jiných ekonomických řad je ta, že hodnoty finanční časové řady jsou zaznamenávány s vyšší frekvencí – denní či dokonce hodinové. Běžné časové řady se sledují většinou v roční, čtvrtletní a měsíční frekvencí. Je tedy zřejmé, že u finančních časových řad odpadá problém datové nedostatečnosti. Tento problém většinou nastává u českých makroekonomických časových řad. České makroekonomické řady jsou totiž obecně „krátké“, v lepším případě obsahují několik desítek pozorování, v horším případě jen několik málo pozorování. Tento fakt je dán tím, že řady jsou vesměs sledovány jen posledních 12 let. Předtím sice byly tyto řady rovněž sledovány, avšak za zcela jiných podmínek (socialistická ekonomika, federace...). Skutečnost, že finanční časové řady jsou sledovány s vyšší frekvencí znamená, že vedle systematických faktorů mají na dynamiku časových řad poměrně značný vliv i faktory nesystematického charakteru. To se projevuje v jejich relativně vysoké a proměnlivé variabilitě.

Mezi typické rysy finančních časových řad patří (viz např. [21])

- *Podmíněná heteroskedasticita*: Finanční časové řady mívají proměnlivou variabilitu. Proto nejlepší výsledky při modelování finančních časových řad dávají modely, které se snaží dynamicky vysvětlit podmíněný rozptyl takových řad (tj. rozptyl podmíněný informací o předchozích hodnotách takové řady). Protože kolísání rozptylu ve statistických modelech se označuje jako heteroskedasticita, používá se v tomto případě termín podmíněná heteroskedasticita.

- *Leptokurtické rozdělení:* V teoretických a empirických pracích zabývajících se finančními časovými řadami se často vychází z toho, že logaritmy výnosů mají normální rozdělení s konstantní střední hodnotou a konstantním rozptylem. Toto rozdělení je symetrické kolem své střední hodnoty, tzn. jeho šikmost je rovna nule. Další vlastností normálního rozdělení je, že koeficient špičatosti je roven třem. V praxi však často dochází k tomu, že rozdělení logaritmů denních, týdenních či čtrnáctidenních výnosů není symetrické a je zešikmené. Tato rozdělení mají „tlustší“ konce než rozdělení normální, tzn. že četnost výskytu extrémně vysokých kladných či záporných výnosů je vyšší než za předpokladu normality. Zároveň se tato rozdělení vyznačují „tenkým pásem“, kdy část hodnot uvažované náhodné veličiny leží s vysokou pravděpodobností blízko její střední hodnotě. Pravděpodobnostní rozdělení takového typu se někdy souhrnně označují jako leptokurtická rozdělení.
- *Výskyt segmentů s nízkou a naopak s vysokou volatilitou:* Typické pro volatilitu finančních časových řad je výskyt segmentů s nízkou a naopak s vysokou volatilitou. Takové segmentování vzniká v důsledku toho, že předchozí vysoká (resp. nízká) volatilita vyvolává s velkou pravděpodobností také vysokou (resp. nízkou) volatilitu v následujícím čase.
- *Stacionarita:* Časové řady měř zisku obvykle fluktuují kolem konstantní (často nulové) úrovně. Také jejich (nepodmíněný) rozptyl lze považovat za konstantní. Toto se označuje jako stacionární chování.
- *Nekorelovanost v časových řadách měř zisku:* Dále se v klasických analýzách finančních časových řad předpokládá, že logaritmy výnosů jsou nekorelované stejně rozdělené náhodné veličiny s nulovou střední hodnotou a konstantním rozptylem (jedná se o proces bílého šumu), nebo nezávislé stejně rozdělené náhodné veličiny s nulovou střední hodnotou a konstantním rozptylem (jedná se o proces striktního bílého šumu). Ve skutečnosti však nemusí být splněna podmínka nekorelovanosti logaritmů výnosů. Nekorelovanost mezi mírami zisku v různých časech však neznamená jejich nezávislost, je to pouze nezávislost lineární, která pak lze vyjádřit stacionárními lineárními modely typu ARMA (viz kapitola 1.2).

- *Významná korelovanost v časových řadách čtverců měr zisku:* Časové řady čtverců měr zisku, na rozdíl od časových řad měr zisku, se obvykle vyznačují významnou korelovaností, proto např. úspěšné modely volatility (kapitola 1.4) konstruují její předpovědi s využitím čtverců měr zisku.
- *Odlišné reakce volatility na velké kladné a na velké záporné hodnoty míry zisku:* Odlišnost v reakci volatility podle toho, zda předchází velká kladná nebo velká záporná hodnota míry zisku, má své ekonomické zdůvodnění a vedla mimo jiné k zobecnění modelů ARCH na GARCH (kapitola 1.4).

### 1.1.3. Akciový trh

Předmětem obchodování na akciovém trhu, jak již bylo výše poznamenáno, jsou akcie. Akcie jsou obchodovatelné cenné papíry, s nimiž jsou spojena práva akcionáře jako společníka podílet se na řízení společnosti (právo účasti a hlasování na valné hromadě akcionářů), na zisku společnosti (právo na dividendy) a na likvidačním zůstatku.

Akcie fungují na rozdíl od dluhopisů jako dividendové cenné papíry (equity), jejichž dividendový výnos není předem zaručen. Dokonce i když je společnost zisková, management může navrhnout zadržení zisku za účelem tvorby fondů pro budoucí investice.

#### 1.1.3.1. Základní pojmy

*Nominální hodnota* akcie představuje podíl na majetku akciové společnosti vyplývající z vlastnictví akcie. Součet nominálních hodnot všech akcií je roven výši základního jmění. *Dividenda* je podíl na zisku společnosti vyplývající z vlastnictví akcie. Výplata dividend není většinou předem zaručena a může mít formu peněžní dividendy, akciové dividendy (akcionář získá nové akcie zdarma či za zvýhodněnou cenu) nebo majetkové dividendy (např. zdarma výrobky či služby související s danou společností). *Tržní cena akcie* (kurs akcie) je cena, za kterou se akcie obchoduje na kapitálovém trhu.

**Volatilita** je míra průměrné intenzity kolísání kurzů cenných papírů a deviz i úrokových sazeb během určitého časového období (obvykle udáváno jako směrodatná odchylka).

Ke každé burze cenných papírů neodmyslitelně patří její **indexy** v roli indikátorů akciového trhu, které koncentrují pohyby cen mnoha akcií do jediného čísla. Definice akciového indexu zní přibližně takto: jedná se o ukazatel (index je latinsky oznamovatel) vývoje kurzů akcií příslušného akciového trhu nebo jen určitého odvětví. Základem indexu jsou akcie vybrané podle předem stanoveného klíče, kterým jsou např. základní jmění obchodované společnosti, její tržní kapitalizace apod., podle čehož se dále odvodí váha vybrané společnosti v sestavovaném indexu.

U specializovaných indexů je důležitá příslušnost k vybranému odvětví nebo regionu. Stanovení hodnoty akciového indexu plyne z poměru mezi jeho aktuální hodnotou a hodnotou indexu při jeho prvním historickém sestavení. Aktuální hodnota indexu se vypočítává v pravidelném intervalu, např. každých 15 vteřin, po celou dobu otevření daného akciového trhu.

Indexy všeobecně nejsou neměnné, je nezbytné je průběžně aktualizovat podle toho, jak se mění situace v ekonomice. Některé společnosti v indexu obsažené zcela zmizí z trhu, jiné se sloučí, zároveň se mění jejich tržní kapitalizace, podíl na výkonnosti ekonomiky a tedy i váha v indexu.

Vývoj akciových indexů je v dlouhodobém výhledu zpravidla růstový, což souvisí s rostoucí výkonností ekonomik a celkové životní úrovně. Existují ovšem i výjimky, kterou je např. japonský index NIKKEI 225, odrážející již deset let trvající hospodářskou krizi v Japonsku. Krátkodobě mohou indexy zaznamenat propady až v řádu desítek procent, které může zapříčinit nenadálá událost, jakou byla např. poměrně nedávná Asijská krize, následující finanční krize v Rusku či teroristický útok na World Trade Center v New Yorku. Kromě denních výkyvů zapříčiněných aktuálním vývojem na trzích dále akciové indexy kolísají ve více méně pravidelných intervalech podle toho, v jaké fázi hospodářského cyklu (růst nebo pokles) se nachází příslušná ekonomika či obor (tzv. cyklické tituly).

Velmi zajímavý vývoj zaznamenaly světové indexy od poloviny 90. let minulého století, kdy pozvolný růst byl nahrazen růstem spíše raketovým. Důvodem byla neuvěřitelně dobrá výkonnost americké ekonomiky, která neměla za posledních několik desetiletí obdoby. Svůj podíl si připsal pochopitelně i rozvoj telekomunikací a Internetu (tzv. New economy) včetně nereálných očekávání z jeho možného ekonomického přínosu. Na burze se tak obchodovaly společnosti s tržní kapitalizací v desítkách milionů dolarů, aniž by přitom vytvořily nějaký zisk.

Z uvedeného vyplývá, že indexy mají opravdu velmi dobrou vypovídací schopnost a investorům se jejich sledování vyplatí. Teoreticky tak mohou včas odhalit případné přehřátí trhu a odchod do bezpečnějších investičních instrumentů, nebo naopak dobu vhodnou k navýšení investic. Skutečnost však bývá trochu jiná kvůli tomu, že každý se snaží vydělat a většinou nevnímá varovné signály.

#### 1.1.3.2. Výpočet indexů

Tržní indexy jsou běžně sestavovány dvěma základními způsoby. Každý z těchto způsobů má svá specifika, a proto je důležité znát způsob jejich tvorby, aby bylo možné korektně interpretovat vývoj těchto indexů resp. stav na tom či onom trhu.

- *Podíl konkrétního titulu na celkovém indexu je dán tržní kapitalizací společnosti.* Tento výpočet vychází z jednoduché logiky, že větší společnosti by měly mít větší podíl na tržním indexu, protože hospodářské výsledky těch největších společností mají mnohem větší vliv na celkovou ekonomiku než činnost menších společností. Tímto způsobem je tvořena většina hlavních světových indexů jako kupříkladu Standard & Poor's 500 Index, NASDAQ Composite Index, Wilshire 5000 Index, London FTSE, indexy MSCI a rovněž i český index PX 50.
- *Cenově vážené indexy.* Konstrukce takových indexů ignoruje celkové tržní kapitalizace společností a spokojuje se pouze se znalostí aktuálních cen jednotlivých akcí. To v praxi znamená, že čím vyšší bude cena akcií společnosti, tím vyšší váhu budou mít tyto akcie v indexu.

Příkladem cenově váženého indexu je jeden z nejznámějších a nejsledovanějších indexů vůbec, americký Dow Jones Industrial Average. Obdobným způsobem je konstruován i japonský Nikkei 225.

Z výše uvedeného je zřejmé, že naprosto stejné události budou ovlivňovat tržní index zcela odlišným způsobem.

### **1.1.3.3. Přehled nejznámějších indexů**

#### **1. Index The Dow Jones Average (DJIA)**

Index The Dow Jones Industrial Average (DJIA) je nejznámějším akciovým indexem světa. Byl vytvořen Charles H. Dowem (respektive Dow Jones Co.) již 26. května 1896. Tehdejších dvanáct amerických průmyslových akcií vytvořilo index s počáteční hodnotou 40,94 bodů. Od roku 1928 má DJIA třicet komponent. Těchto 30 amerických průmyslových, mediálních, finančních a technologických blue-chips v současné době reprezentuje asi pětinu tržní kapitalizace všech amerických společností a asi čtvrtinu kapitalizace Newyorské akciové burzy (NYSE). Typy společností, které jsou zahrnuty do tohoto indexu, jsou tedy různorodé a pokrývají veškeré hlavní oblasti americké ekonomiky, kromě dopravy a společností veřejně prospěšných. Výběr společností, které mají být zařazeny do báze DJIA, je na úvaze redaktorů časopisu The Wall Street Journal a je v případě potřeby revidován. Původní DJIA byl jednoduše průměr cen akcií. Dnes se užívá cenový vážený systém. Výhodou tohoto indexu je, že obstál ve zkoušce času. Další výhodou může být to, že obsahuje 30 nejvýznamnějších společností v USA a není považován za volativní (nestálý) nebo riskantní. Nevýhodou je, že existuje více než 10 000 obchodních společností v USA a DJIA obsahuje pouze 30 společností. Z tohoto důvodu Standard&Poor's 500 index (S&P 500) začíná přebírat jeho funkci. Kromě toho vážení na základě kapitalizace trhu (tak je počítán Standard&Poor's 500 index) je obecně považováno za efektivnější než vážení na základě ceny.

#### **2. Index The Standard & Poor's 500**

Jak již bylo výše řečeno, hlavní nevýhodou DJIA je, že obsahuje pouze 30 společností. Index Standard&Poor's 500 je v tomto ohledu vylepšením DJIA, protože zahrnuje 500 společností a navíc se na něj pohlíží jako na základní kritérium amerického

akciového trhu. Společnosti, které jsou zahrnuty do báze tohoto indexu, pokrývají veškeré hlavní oblasti americké ekonomiky. Nejedná se o 500 největších společností, ale spíše o 500 společností s největším dosahem na trhu vybraných s ohledem na velikost trhu, likviditu a průmyslový sektor. Společnosti zahrnuté do báze indexu S&P 500 vybírá výbor S&P indexu. Každý rok se provádí 25–50 změn v bázi, a to z důvodů vytvoření fúzí nebo úpadků některých společností. V minulosti byly zahrnuty i mezinárodní společnosti, ale v budoucnosti budou přidávány pouze americké společnosti. Index S&P 500 je index vážené tržní kapitalizace, tzn. že každá akcie v indexu je zastoupena v poměru ke své tržní kapitalizaci. Výhodou tohoto indexu je, jak již bylo uvedeno, že zahrnuje 500 společností, tzn. nabízí velkou diverzifikaci a zahrnuje přibližně 70% amerického trhu. Naopak nevýhodou je, že 45 špičkových společností zahrnutých do indexu obsahuje více než 50% hodnoty indexu. Další nevýhodou je, že je v něm obsažen velmi malý počet zahraničních společností.

### 3. Dow Jones EURO STOXX 50 Index

Tento index je jedním z hlavních indexů v Evropě. Zahrnuje 50 evropských blue-chips (nejlepších) akcií zemí eurozóny. Ty dohromady tvoří téměř 60 % souhrnné tržní kapitalizace zdejších akciových trhů. Index obsahuje například akcie společností Nokia, Unilever, Siemens, Deutsche Bank, Philips, Allianz, Danone, Ahold nebo Volkswagen. Vznikl v únoru 1998 s bází vztaženou k 31. prosince 1991 při zahajovací hodnotě 1000 bodů. Došlo vlastně ke spojení německé, švýcarské a pařížské burzy. Index obsahuje 4 základní indexy a 19 oborových indexů počítaných pro západní Evropu a země participujících v eurozóně. Jsou to: Dow Jones Stoxx (všeobecný evropský index), Dow Jones Stoxx 50 (evropský blue-chip index –50 blue chips akcií získaných z Dow Jones Stoxx), Dow Jones Euro Stoxx (všeobecný Euro index 2, Dow Jones Stoxx kromě zemí v EMU), Dow Jones Euro Stoxx (50 Euro blue-chips akcií, získaných z Dow Jones Euro Stoxx) a 19 oborových indexů získaných z Dow Jones Stoxx a Dow Jones Euro Stoxx. Dow Jones Stoxx, Dow Jones Euro Stoxx a příslušné oborové indexy jsou měřítkem výkonnosti akciového trhu, zatímco zbývající dva indexy jsou zaměřeny na derivátové produkty. V lednu 1999 začalo 11 evropských zemí obchodovat na akciovém trhu Evropy

v měně euro Tento index reprezentuje akciový trh Evropy a hlavní burza se nachází ve švýcarském Curychu. Index je indexem vážené tržní kapitalizace.

#### 4. Index PX 50 a jeho nástupce PX

Česká burza zavedla svůj oficiální index PX 50 při příležitosti prvního výročí zahájení obchodování. Byl zvolen standardní výpočet indexu ve shodě s metodologií IFC (International Finance Corporation) doporučenou pro tvorbu indexů na vznikajících trzích. Na základě rozborů bylo rozhodnuto vytvořit bázi složenou z 50 emisí. Počet bazických emisí byl variabilní. V souladu se Zásadami aktualizace báze indexu PX 50 schválenými v prosinci 2001 však nemohl převýšit padesát. Do báze indexu se nezařazovaly emise oboru č.18 (investiční fondy) a holdingových společností vzniklých transformací z investičních fondů, neboť v jejich kursech se již promítaly cenové pohyby bazických emisí.

PX 50 Index byl váženým indexem padesáti hlavních titulů pražské burzy. V následující tabulce (tabulka je aktuální k únoru 2006) jsou uvedeny společnosti, které se výrazně podílely na tvorbě indexu.

*Tabulka 1 – Seznam společností podílejících se na tvorbě indexu PX 50 k únoru 2006*

Pořadí	Název CP	Redukovaný počet CP	Redukovaná tržní kapitalizace (mil.Kč)	Váha (%)
1	ČEZ	207 851 484	172 724,6	21,60
2	ERSTE BANK	115 923 932	155 569,9	19,45
3	ČESKÝ TELEKOM	293 990 325	153 252,7	19,16
4	KOMERČNÍ BANKA	38 009 852	129 499,6	16,19
5	UNIPETROL	181 334 764	53 548,2	6,7
6	ZENTIVA	38 136 230	45 229,6	5,66
7	CETV	30 481 734	42 491,5	5,31
8	PHILIP MORRIS ČR	1 913 698	35 122,1	4,39
9	ORCO	5 535 224	12 326,9	1,54
<b>Celkem</b>			799 769,5	100



Za výchozí burzovní den byl zvolen 5. 4. 1994, výchozí hodnotou indexu PX 50 se stalo 1 000 bodů.

Index PX 50 byl počítán na základě tržní kapitalizace. Vzorec pro výpočet indexu PX 50 byl následující:

$$PX(t) = K(t) \cdot \frac{M(t)}{M(0)} \cdot 1000 \quad (1.1.5)$$

kde

$M(t)$  je tržní kapitalizace báze v čase  $t$ ,

$M(0)$  je tržní kapitalizace v základním (výchozím) období,

$K(t)$  je faktor zřetězení v čase  $t$  (zohledňuje změny provedené v bázi indexu).

Výpočetní vzorec (1.1.5) lze přepsat do tvaru:

$$PX(t) = \frac{\text{kapitalizace báze v čase } t}{\text{hodnota báze v čase } t} \cdot 1000 \quad (1.1.6)$$

Z porovnání vzorců (1.1.5) a (1.1.6) je zřejmé, že hodnota báze v čase  $t$  uvedená ve jmenovateli předchozího zlomku je rovna výrazu  $M(0)/K(t)$ , tedy korigované tržní kapitalizaci báze v základním období.

Ke změně hodnoty báze ve vztahu (1.1.6) docházelo nejen při výměně bazické emise, ale i při všech operacích, které mění tržní kapitalizaci bazické emise (zvýšení, resp. snížení počtu cenných papírů (CP) v emisi, splnutí emisí, emise předkupních práv). Speciální operace výplata dividendy nezpůsobovala v případě indexu PX 50 změnu hodnoty báze, dividendové výnosy se nezohledňovali. Jednalo se tedy o **index cenový**, nikoliv o "return" index. Platné počty CP v bazických emisích vycházely z údajů poskytovaných Střediskem cenných papírů (SCP). Transformace hodnoty báze byla založena na principu spojitosti indexu v okamžiku změny báze, z něhož plyne následující vztah:

$$\text{Nová hodnota báze} = \text{stará hodnota báze} \cdot \frac{\text{kapitalizace nové báze v čase } t}{\text{kapitalizace staré báze v čase } t} \quad (1.1.7.)$$

Indexy pražské burzy PX 50 a PX-D byly 20.března 2006 nahrazeny jediným indexem PX. Výpočet obou indexů byl ukončen 17.března 2006. Cílem bylo vytvořit index, který by lépe odpovídal požadavkům připravovaného derivátového trhu. Nový index může sloužit jako podkladové aktivum pro obchodování s futures na index a zároveň se vyřešil problém s korelací stávajících indexů, které se vyvíjely téměř totožně. Index PX si uchoval všechny důležité vlastnosti svých předchůdců. Těmi jsou v případě indexu PX 50 jeho velmi cenná, více než dvanáctiletá historie a v případě indexu PX-D, z pohledu likvidity, optimální složení báze. Pravidla výběru bazických emisí nového indexu totiž zabezpečují, že do báze může být zařazena pouze likvidní emise. Výchozí báze tak obsahuje emise obchodované v segmentu SPAD, což je charakteristická vlastnost báze indexu PX-D. Index PX tedy spojitě navázal na vývoj indexu PX 50 a je rovnocenným nástupcem obou stávajících indexů, přičemž slučuje v sobě jejich přednosti.

Hodnota indexu PX-D se po 20. březnu 2006 stanovuje za pomoci tzv. PX-D koeficientu vynásobeného hodnotou indexu PX. Koeficient pro tuto operaci vznikl podílem závěrečných hodnot indexů PX-D a PX 50 ze dne 17.3. 2006 a jeho hodnota byla burzou zveřejněna téhož dne.

#### 1.1.3.4. Akciové analýzy

Klíč k úspěšnému obchodování na akciovém trhu je velmi jasný a jednoduchý: ve správný okamžik akcie levně koupit a ve správný okamžik je draze se ziskem prodat. Problémem je, že kdy nastane ten správný okamžik nikdo neví. Proto vznikly akciové analýzy, které se na základě analýzy různých faktorů snaží nalézt ty, které tržní cenu akcií nějakým způsobem ovlivňují. Základní informace o akciových analýzách můžeme najít např. v [70].

- **Fundamentální analýza**

Fundamentální analýza se snaží předpovědět jaké faktory působí a ovlivňují pohyby kurzu jednotlivých akcií. Postup prognózy se děje ve třech krocích, a to globální analýza, odvětvová analýza a analýza konkrétní akcie. Globální analýza zkoumá krátkodobé i dlouhodobé vlivy ekonomických makroagregátů na ceny akcií (inflace, hospodářského růst, úrokových sazeb atd.). Odvětvová analýza měří citlivost odvětví na hospodářský

cyklus, rozsah a způsob vládní regulace, sílu odborů, míru inovací v daném odvětví atd. Pomocí finanční analýzy jednotlivých titulů se pak stanoví odhad vnitřní hodnoty příslušné akcie.

- **Psychologická analýza**

Psychologická analýza je založena na předpokladu, že na kurz akcií má velký vliv psychologická reakce investorů zejména v krátkodobém časovém horizontu. Zakládá se na předpokladu, že v ekonomii se nedá vše spočítat a reakce lidí nejsou založeny na racionálních úvahách. Proto se snaží odhalovat změny v psychologickém chování investorů.

- **Technická analýza**

Technická analýza je metoda odhadu cen akcií, založená na studiu chování jednotlivých akcií a celkového trhu. Pravděpodobný budoucí vývoj cen se předpovídá na základě vývoje kurzů a objemů obchodů v minulosti ( máme tedy k dispozici časové řady tržních cen, objemů obchodů a různých indexů). Přitom se předpokládá, že

1. Tržní ceny akcií obsahují veškeré relevantní ekonomické, politické a psychologické informace. Skutečná cena akcie je určena poměrem nabídky a poptávky.
2. Cenový pohyb vždy vytváří určitý trend, ve kterém cena po jistou dobu setrvává.
3. Vzory tržního chování investorů se opakují a proto lze z minulého vývoje usuzovat na budoucí vývoj.

Technický analytik se pokouší rozpoznat v pohybu kursu určitý tvar (formaci) a podle něj časuje nákup a prodej libovolného cenného papíru (timing). K tomu používá jednak grafické metody (např. linie podpory a odporu) a jednak technické indikátory. Technický indikátor je funkce, která pro každý den  $t$  nabývá určité reálné hodnoty, závislé na minulých cenách a objemech konkrétní akcie. Technické indikátory jsou součástí obchodních systémů a generují signály k nákupu (BUY), k prodeji (SELL) anebo k držení (HOLD) akcií. Používané indikátory jsou např. momentum, index relativní síly, linie růstu a poklesu (*AD-Linie*), míra změny atd. (viz [83]).

- **Teorie efektivních trhů**

Teorie efektivních trhů (*Efficient market*) je v současnosti nejvýznamnějším proudem v ekonomii. Teorie tvrdí, že akciový trh je efektivní a kurzotvorné informace v sobě obsahuje téměř okamžitě (rozšíření důležité informace trvá maximálně 30 vteřin). Proto selhávají jednotlivé obchodní strategie, protože kurz je vždy objektivní a přizpůsobuje se své vnitřní hodnotě a na pohyby kurzu má vliv pouze neočekávaná informace a proto se kurzy pohybují též neočekávaně.

Předpoklady efektivního trhu jsou:

1. Na akciovém trhu působí obrovské množství racionálních investorů, kteří průběžně analyzují i obchodují.
2. Investoři mají k dispozici dostatek pravdivých, levných a aktuálních informací.
3. Investoři reagují přesně a rychle na nové informace.
4. Transakční náklady jsou nízké.
5. Trh je vysoce likvidní.
6. Žádný účastník nemá monopolní či výsadní postavení.
7. Kvalitní infrastruktura a právní regulace trhu.

I přestože teorie efektivních trhů předpokládá, že kurz je vždy objektivní, dochází k pravidelně opakujícím se anomáliím (akvizice a fúze, kótování na burzách, efekt malých společností, efekt nízkého P/E, pondělní efekt, lednový efekt), které však lze víceméně spolehlivě vysvětlit.

## 1.2. Box- Jenkinsova metodologie

Nejjednodušší modely popisující chování finančních časových řad jsou lineárně stochastické modely. Základní principy stochastické koncepce tvorby lineárních modelů vycházejí z Box – Jenkinsovy metodologie, jejíž základy jsou uvedeny v publikaci Boxe a Jenkinse ([14]).

### 1.2.1. Box -Jenkinsova metodologie : Stacionární procesy

V Box – Jenkinsově metodologie se za základní prvek konstrukce modelu časové řady bere reziduální složka, která může být tvořena korelovanými (závislými) náhodnými veličinami. Box - Jenkinsova metodologie může tedy zpracovávat časové řady s navzájem závislými pozorováními. Těžiště jejich postupů tudíž spočívá ve využití korelační analýzy. Analýza časové řady se v rámci této metodologie provádí systematicky podle předem daného klíče:

- identifikace modelu,
- odhad parametrů modelu,
- ověřování modelu.

#### 1.2.1.1. Základní pojmy a aparát Box -Jenkinsovy metodologie

Jednou z nejdůležitějších vlastností časové řady je její **stacionarita**. Stacionarita časové řady znamená, že chování řady je v jistém smyslu stochasticky ustálené.

Rozlišujeme

1. *Striktní stacionaritu* – pravděpodobnostní chování příslušného stochastického procesu je invariantní vůči posunům v čase, tzn. pravděpodobnostní chování náhodného vektoru  $(y_{t1}, y_{t2}, \dots, y_{tk})$  je stejné jako rozdělení vektoru  $(y_{t1+h}, y_{t2+h}, \dots, y_{tk+h})$  pro libovolné  $h$ .
2. *Slabou stacionaritu* – stochastický proces má konstantní úroveň (tj. konstantní střední hodnotu), je rovnoměrně vyvážený kolem své konstantní úrovně (tj. má konstantní rozptyl) a kovarianční strukturu druhého řádu je invariantní vůči posunům v čase, tj.

$$\text{cov}(y_t, y_s) = \text{cov}(y_{t+h}, y_{s+h}) \text{ pro libovolné } h. \quad (1.2.1)$$

Závislost mezi dvěma libovolnými pozorováními spočívá pouze v jejich vzájemné časové vzdálenosti a ne na jejich skutečném časovém umístění v řadě.

Slabá a silná stacionarita splývá v případě normálního procesu. Pokud proces není normální většinou se předpokládá slabá stacionarita, neboť je méně omezující.

Základním prostředkem podávajícím informaci o charakteru stochastického procesu je **autokorelační funkce (ACF)** a **parciální autokorelační funkce (PACF)**. Na základě chování těchto funkcí lze identifikovat model konkrétních analyzovaných časových řad.

Vzhledem k tomu, že autokorelační funkce je definována pomocí autokovarianční funkce, definujeme nejdříve autokovarianční funkci. Pro stacionární (slabě) časovou řadu  $y_t$  se hodnota její *autokovarianční funkce*  $\gamma_k$  v bodě  $k$  definuje jako

$$\gamma_k = \text{cov}(y_t, y_{t+h}) = E(y_t - \mu)(y_{t+h} - \mu), \quad k = \dots -1, 0, 1, \dots \quad (1.2.2.)$$

Hodnota *autokorelační funkce*  $\rho_k$  v bodě  $k$  se definuje jako

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\gamma_k}{\sigma_y^2}, \quad k = \dots -1, 0, 1, \dots, \quad (1.2.3.)$$

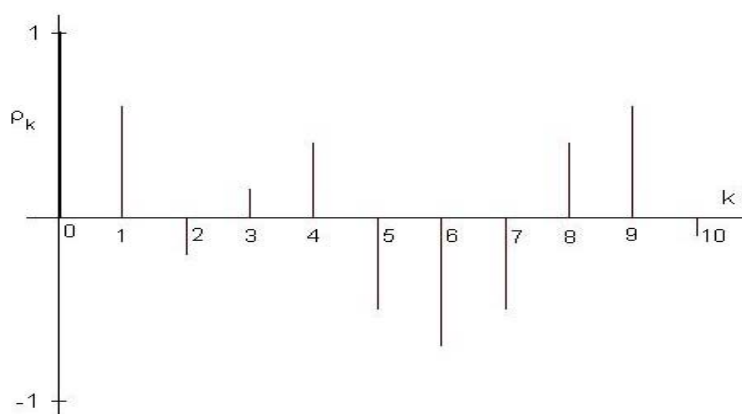
kde  $\mu = Ey_t$  je střední hodnota a  $\sigma_y^2 = \gamma_0 = \text{var}(y_t)$  je rozptyl dané stacionární řady.

V případě stacionárního stochastického procesu má autokovarianční a autokorelační funkce následující vlastnosti:

1.  $\rho_0 = 1$ ,
2.  $|\gamma_k| \leq \gamma_0$ ;  $|\rho_k| \leq 1$ ,
3.  $\gamma_k = \gamma_{-k}$   
 $\rho_k = \rho_{-k}$  pro všechna  $k$ .

Vlastnost 3 znamená, že autokovarianční a autokorelační funkce jsou sudé. Ze sudosti těchto funkcí vyplývá, že se stačí omezit na  $k \geq 0$ .

Grafický záznam  $\rho_k$  pro jednotlivá  $k \geq 0$  se nazývá **korelogram**, jednu z jeho forem ukazuje obr.1.



Obr.1: Korelogram

**Parciální autokorelační koeficient**  $\rho_{kk}$  je dalším důležitým pojmem Box – Jenkinsovy metodologie. Podává informaci o korelaci veličin  $y_t$  a  $y_{t-k}$  očištěnou o vliv veličin ležících mezi nimi  $y_{t-1}, y_{t-2}, \dots, y_{t-k+1}$ . Korelace mezi dvěma náhodnými veličinami je totiž často způsobena tím, že obě tyto veličiny jsou korelovány s veličinou třetí nebo více veličinami.

Parciální autokorelační funkce  $\rho_{kk}$  je definována jako

$$\rho_{kk} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_{k-2} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & 1 \end{vmatrix}}. \quad (1.2.4)$$

Autokorelační (ACF) a parciální autokorelační (PACF) funkce dané řady se v bodě  $k=1$  shodují. Je tedy

$$\rho_{11} = \rho_1 \quad (1.2.5)$$

Pro  $k = 2$  platí

$$\rho_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}. \quad (1.2.6)$$

Při empirické práci s časovými řadami pracujeme ovšem s odhady autokovarianční funkce, autokorelační funkce a parciální autokorelační funkce.

Uvažujme-li napozorovanou řadu  $y_1, y_2, \dots, y_n$ . Potom *odhad střední hodnoty* je

$$\bar{y} = \frac{1}{n} \sum_{t=1}^n y_t, \quad (1.2.7)$$

*odhad autokovarianční funkce* je

$$c_k = \sum_{t=1}^{n-k} \frac{(y_t - \bar{y})(y_{t+k} - \bar{y})}{n}, \quad k = 0, 1, \dots, n-1, \quad (1.2.8)$$

*odhad autokorelační funkce* je

$$r_k = \frac{c_k}{c_0}, \quad k = 0, 1, \dots, n-1. \quad (1.2.9)$$

Aby odhady měly praktický význam požaduje se obvykle  $n > 50, k < n/4$ .

*Odhady  $r_{kk}$  parciální autokorelační funkce  $\rho_{kk}$  se počítají rekurentně podle následujících vzorců*

$$r_{11} = r_1$$

$$r_{kk} = \frac{r_k - \sum_{j=1}^{k-1} r_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j} r_j} \quad \text{pro } k > 1, \quad (1.2.10)$$

kde

$$r_{kj} = r_{k-1,j} - r_{kk} r_{k-1,k-j} \quad \text{pro } j=1, 2, \dots, k-1. \quad (1.2.11)$$

Jak už bylo řečeno, chování autokorelační a parciální autokorelační funkce jsou důležitým ukazatelem, napovídají jaký typ modelu je vhodné pro danou řadu použít. Pro



identifikaci modelu je důležité určit hodnotu  $k_0$  (identifikační bod), za kterou začínají být tyto funkce nulové. Vzhledem k tomu, že místo teoretické autokorelační a parciální autokorelační funkce uvažujeme jejich odhady, bude nás zajímat, jak blízko nule musí být tyto odhady, abychom s předem danou spolehlivostí mohli tvrdit, že  $\rho_k = 0$ , popř.  $\rho_{kk} = 0$ .

V případě autokorelační funkce  $\rho_k$  o nulovosti rozhodneme na základě *Bartlettovy aproximace* ([9]): je-li  $\rho_k = 0$  pro  $k > k_0$ , pak

$$\sigma(r_k) = \sqrt{\text{var}(r_k)} \approx \sqrt{\frac{1}{n} \left( 1 + 2 \sum_{j=1}^{k_0} r_j^2 \right)}, \quad k > k_0 \quad (1.2.12)$$

( $\sigma(r_k)$  je směrodatná odchylka odhadu  $r_k$  autokorelační funkce  $\rho_k$ ).

Testujeme tedy

$$\begin{aligned} H_0: \rho_k &= 0 \\ \text{proti} & \\ H_1: \rho_k &\neq 0 \end{aligned} \quad (1.2.13)$$

Je-li

$$|r_k| > 2 \cdot \sigma_k, \quad (1.2.14)$$

hypotézu  $H_0$  zamítneme na hladině spolehlivosti 0,05.

V případě parciální autokorelační funkce o nulovosti rozhodneme na základě *Quenouilleovy aproximace* ([69]) pro směrodatnou odchylku odhadu  $r_{kk}$ : je-li  $\rho_{kk} = 0$  pro  $k > k_0$ , pak

$$\sigma(r_{kk}) = \sqrt{\text{var}(r_{kk})} \approx \sqrt{\frac{1}{n}}, \quad k > k_0. \quad (1.2.15)$$

(Předpokládá se, že uvažovaný proces vyhovuje tzv. autoregresnímu modelu řádu  $k_0$ .)

Testujeme

$$H_0: \rho_{kk} = 0$$

proti (1.2.16)

$$H_1: \rho_{kk} \neq 0.$$

Je-li

$$|r_{kk}| > 2 \cdot \sigma(r_{kk}), \quad (1.2.17)$$

hypotézu zamítneme na hladině spolehlivosti 0,05.

### 1.2.1.2. Lineární proces

Lineárním procesem nazýváme řadu tvaru

$$y_t = \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots, \quad (1.2.18)$$

kde  $\varepsilon_t$  je bílý šum s rozptylem  $\sigma_\varepsilon^2$  a  $\psi_j$  jsou parametry.

Často se využívá operátor zpětného posunutí, který je definován jako

$$B y_t = y_{t-1}. \quad (1.2.19)$$

Při opakovaném použití platí

$$B^j y_t = y_{t-j}. \quad (1.2.20)$$

Lineární proces potom můžeme zapsat ve tvaru

$$y_t = \psi(B) \varepsilon_t, \quad (1.2.21)$$

kde

$$\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots = 1 + \sum_{j=1}^{\infty} \psi_j B^j. \quad (1.2.22)$$

Postačující podmínka pro existenci lineárního procesu (1.2.18) je, že řada v (1.2.22) konverguje podle kvadratického středu, tj.  $\psi(B)$  konverguje pro  $|B| \leq 1$ .

Pokud současná hodnota  $y_t$  lineárního procesu jde vyjádřena pomocí jeho minulých hodnot a současné hodnoty bílého šumu, tj.

$$y_t = \pi_1 y_{t-1} + \pi_2 y_{t-2} + \dots + \varepsilon_t, \quad (1.2.23)$$

nazývá se tento proces invertibilní.

S využitím operátoru zpětného posunutí lze ekvivalentně psát

$$\pi(B)y_t = \varepsilon_t, \quad \text{kde} \quad \pi(B) = 1 - \pi_1 B - \pi_2 B^2 - \dots = 1 - \sum_{j=1}^{\infty} \pi_j B^j. \quad (1.2.24)$$

Postačující podmínka pro invertibilitu lineárního procesu ([14]) je

$$\pi(B) \text{ konverguje pro } |B| \leq 1.$$

Vztah mezi  $\pi_j$  a  $\psi_j$  lze schematicky zapsat jako

$$\psi(B)\pi(B) = 1. \quad (1.2.25)$$

Vynulováním všech parametrů lineárního procesu až na konečný počet dostaneme procesy, které mají praktický význam. Jsou to: *proces klouzavých součtů řádů  $q$  (MA( $q$ )), autoregresní proces řádů  $p$  (AR( $p$ )) a smíšený proces řádů  $p$  a  $q$  (ARMA( $p, q$ )).*

#### 1.2.1.2.1. Proces klouzavých součtů MA( $q$ )

Proces klouzavých součtů řádů  $q$  značený jako MA( $q$ ) má následující tvar

$$y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}. \quad (1.2.26)$$

Pokud použijeme symbolu zpětného posunutí, dá se zapsat jako

$$y_t = \theta(B)\varepsilon_t, \quad \text{kde} \quad \theta(B) = 1 + \sum_{j=1}^q \theta_j B^j. \quad (1.2.27)$$

Polynom  $\theta(B)$  je tzv. operátor klouzavých součtů,  $\varepsilon_t$  je bílý šum a  $\theta_j$  jsou parametry.

Proces MA( $q$ ) má následující vlastnosti:

- je stacionární pro libovolnou volbu jeho parametrů  $\theta_j$ ,
- proces je invertibilní, jestliže všechny kořeny polynomu  $\theta(B)$  leží vně jednotkového kruhu (v komplexní rovině),
- střední hodnota je nulová,
- rozptyl  $\sigma_y^2$  je roven  $\sigma_y^2 = (1 + \theta_1^2 + \dots + \theta_q^2) \sigma_\varepsilon^2$ , (1.2.28)
- autokorelační funkce  $\rho_k$  má tvar

$$\rho_k = \frac{\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q}{1 + \theta_1^2 + \dots + \theta_q^2} \quad \text{pro } k = 1, 2, \dots, q \quad (1.2.29.)$$

$$\rho_k = 0 \quad \text{pro } k > q.$$

Identifikační bod  $k_0$  pro autokorelační funkci je  $k_0 = q$ .

- parciální autokorelační funkce je omezena buď geometricky klesající posloupností nebo sinusoidou s geometricky klesající amplitudou.

Nejpoužívanější modely procesu klouzavých součtů jsou 1. a 2.řádu (MA(1), MA(2)).

#### 1.2.1.2.2. Autoregresní proces AR(p)

Autoregresní proces řádu  $p$  značený jako AR(p) je definován jako

$$y_t = \varphi_1 y_{t-1} + \dots + \varphi_p y_{t-p} + \varepsilon_t. \quad (1.2.30)$$

Pomocí symbolu zpětného posunutí  $B$  se dá zapsat jako

$$\varphi(B)y_t = \varepsilon_t, \quad (1.2.31)$$

kde  $\varphi(B) = 1 - \sum_{j=1}^p \varphi_j B^j$  je tzv. autoregresní operátor.

Proces AR(p) má následující vlastnosti:

- proces je vždy invertibilní (pro libovolnou volbu parametrů  $\varphi_j$ ),
- proces je stacionární, jestliže všechny kořeny polynomu  $\varphi(B)$  leží vně jednotkového kruhu (v komplexní rovině),
- střední hodnota procesu je nulová,
- rozptyl  $\sigma_y^2$  je roven  $\sigma_y^2 = \frac{\sigma_\varepsilon^2}{1 - \varphi_1 \rho_1 - \dots - \varphi_p \rho_p}$ , (1.2.32)
- autokorelační funkce je lineární kombinací klesajících geometrických posloupností a sinusoid různých frekvencí s geometricky klesajícími amplitudami,
- pro parciální autokorelační funkci procesu AR(p) je identifikační bod  $k_0 = p$ , tj.  $\rho_{kk} = 0$  pro  $k > p$ .

Podobně jako u procesu klouzavých součtů se v praxi používají modely nejvýše řádu druhého.

### 1.2.1.2.3. Smíšený proces ARMA(p,q)

Smíšený proces řádu p a q značený jako ARMA(p,q) je definován jako

$$y_t = \varphi_1 y_{t-1} + \dots + \varphi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}. \quad (1.2.33)$$

Pokud použijeme symbolu zpětného posunutí B, dá se zapsat jako

$$\varphi(B)y_t = \theta(B)\varepsilon_t, \quad (1.2.34)$$

$$\text{kde } \varphi(B) = 1 - \sum_{j=1}^p \varphi_j B^j, \quad \theta(B) = 1 + \sum_{j=1}^q \theta_j B^j. \quad (1.2.35)$$

Proces ARMA(p,q) má následující vlastnosti:

- podmínka stacionarity je totožná s podmínkou stacionarity procesu AR(p) a podmínka invertibility je totožná s podmínkou invertibility procesu MA(q),
- střední hodnota procesu je nulová,
- autokorelační funkce je ve tvaru lineární kombinace klesajících geometrických posloupností nebo sinusoid s geometricky klesající amplitudou po prvních  $q-p$  hodnotách,
- parciální autokorelační funkce je ve tvaru lineární kombinace klesajících geometrických posloupností nebo sinusoid s geometricky klesající amplitudou po prvních  $p-q$  hodnotách.

Při praktických aplikacích Box – Jenkinsovy metodologie se obvykle vystačí s procesy, pro něž  $p + q \leq 2$ , tj.s autoregresními procesy do maximálně druhého řádu a procesy klouzavých součtů též do maximálně druhého řádu.

### 1.2.1.3. Výstavba modelu ARMA(p,q)

Výstavbu modelu lze rozdělit do tří základních kroků

- identifikace modelu,
- odhad parametrů modelu,
- ověřování modelu.

### Identifikace modelu

Identifikace modelu je první fází výstavby modelu. Jejím úkolem je rozhodnout, jaký typ modelu vybrat, tj. zda AR, MA nebo ARMA, a explicitně určit řád modelu. Před vlastní identifikací se obvykle pořídí grafický záznam řady, pomocí něhož se opticky zkontroluje stacionarita dané řady, tj. zda charakter řady zůstává v čase stejný včetně udržování konstantní úrovně. V případě, že střední hodnota stacionárního procesu je nenulová, je nutné danou řadu centrovat. Po těchto přípravných operacích lze přistoupit k vlastní identifikaci, která je založena na zkoumání průběhu odhadnuté autokorelační funkce  $r_k$  a parciální autokorelační funkce  $r_{kk}$ . Stačí obvykle použít prvních 20 hodnot těchto funkcí. Snažíme se především zjistit existenci případného identifikačního bodu  $k_0$  (testy byly uvedeny v kapitole 1.2.1.1). Vzhledem k tomu, že pracujeme s odhady autokorelační a parciální autokorelační funkce, je vhodné netrvat na jednom modelu, ale vzít v úvahu a přezkoušet několik alternativ.

### Odhad parametrů modelu

Po identifikaci modelu lze přistoupit k druhé fázi výstavby modelu, a to odhadu parametrů modelu. Pokud se rozhodneme pro nějaký typ modelu, určíme počáteční odhady parametrů podle vzorců uvedených v [8]. Tyto hrubé odhady je možné použít jako počáteční hodnoty v iteračních odhadových procedurách. Popis těchto procedur je značně komplikovaný a v současné době existují statistické pakety, které požadovaný odhad parametrů provedou. Odhadování parametrů v modelech časových řad můžeme provádět např. podmíněnou metodou nejmenších nelineárních čtverců nebo nepodmíněnou metodou nejmenších nelineárních čtverců. Ve studii [8] bylo ukázáno, že pro řady s délkou v průměru větší než 75 je rozdíl mezi oběma metodami zanedbatelný, takže vzhledem k výpočetní jednoduchosti se pak preferuje podmíněný přístup. V sezónních modelech se naopak doporučuje používat vždy přístup nepodmíněný.

### Ověřování modelu

Třetí fáze výstavby modelu v Box – Jenkinsově metodologii je ověřování modelu zkonstruovaného v předchozích dvou fázích. Toto ověřování má potvrdit nebo zamítnout adekvátnost modelu. Jestliže jsou zjištěny nesrovnalosti mezi danou časovou řadou a zkonstruovaným modelem, je třeba celý třífázový postup zopakovat. V tomto případě by však

zároveň měly být v rámci ověřování modelu získány informace, jakým způsobem model opravit. Vzhledem k tomu, že metody posuzují adekvátnost modelu z různých hledisek a s různou účinností, doporučuje se použít vždy více takových metod současně. Mezi metody ověřování adekvátnosti modelu patří např. metoda přeparametrizování modelu, metoda odhadnutých reziduí nebo Portmanteau test. Výše uvedená kritéria můžeme najít v publikaci [20].

Při analýze časové řady můžeme dojít k závěru, že existuje několik akceptovatelných modelů. Někdy je jednoduché z této množiny vybrat ten nejlepší. Může ale nastat situace, kdy výběr nejlepšího modelu je velice obtížný. Pro řešení této úlohy bylo navrženo několik dodatečných kritérií. Tyto kritéria jsou založena na porovnání reziduí jednotlivých modelů prostřednictvím souhrnných statistik. Mezi nejpoužívanější patří (viz např.[17])

- **Akaikeho kritérium AIC**

Akaikeho kritérium AIC má následující tvar

$$AIC(M) = n \ln s^2 + 2M, \quad (1.2.36)$$

kde  $s^2$  je reziduální rozptyl použitého modelu ARMA(p,q),  $n$  je počet pozorování a  $M = p + q$  je celkový počet parametrů tohoto modelu .

- **Akaikeho kritérium BIC**

Bylo dokázáno, že kritérium AIC vede k nadhodnocení řádu autoregrese, tudíž Akaike (1979) toto kritérium rozšířil do formy

$$BIC(M) = n \ln s^2 - (n - M) \ln \left( 1 - \frac{M}{n} \right) + M \ln n + M \ln \left[ \frac{1}{M} \left( \frac{s_y^2}{s^2} - 1 \right) \right], \quad (1.2.37)$$

kde  $s^2$  je reziduální rozptyl použitého modelu ARMA(p,q),  $n$  je počet pozorování,  $M = p + q$  je celkový počet parametrů tohoto modelu a  $s_y^2$  značí výběrový rozptyl zkoumané časové řady.

- **Schwartzovo bayesovské kritérium SBC**

Schwartzovo bayesovské kritérium SBC má tvar

$$SBC(M) = n \ln s^2 + M \ln n, \quad (1.2.38)$$

kde  $s^2$  je reziduální rozptyl použitého modelu ARMA(p,q),  $n$  je počet pozorování a  $M = p + q$  je celkový počet parametrů tohoto modelu.

Ve všech případech se vybírá model, který minimalizuje hodnotu použitého kritéria.

## 1.2.2. Box – Jenkinsova metodologie: Nestacionární procesy

Původní finanční časové řady (kurzy akcií) mají zpravidla nestacionární charakter, ať už z hlediska časově proměnné střední hodnoty nebo časově proměnného rozptylu. Tyto řady lze v některých případech modelovat pomocí modelů ARIMA. Tyto modely totiž modelují stochasticky vedle náhodných fluktuací i trendovou složku. Stacionarity lze v mnohých případech dosáhnout diferencováním původní časové řady. Pro časové řady z praxe se přitom málokdy použije řád diferencování větší než 2. Určení řádu diferencování můžeme najít např. v [20]. Diferencováním se tedy nestacionární řada převede na stacionární a poté lze proces vyjádřit pomocí modelů ARMA.

### 1.2.2.1. Integrovaný smíšený model ARIMA (p,d,q)

Smíšený model ARIMA (p,d,q) definujeme jako

$$\varphi(B)w_t = \theta(B)\varepsilon_t, \quad (1.2.39)$$

kde

$$w_t = \Delta^d y_t \quad (1.2.40)$$

je d-tá diference modelovaného procesu  $y_t$ . Model (1.2.39) je stacionární model ARMA(p,q) pro proces  $w_t$ .

Je zřejmé, že diferenční operátor  $\Delta$  lze vyjádřit pomocí operátoru zpětného posunutí, neboť

$$\begin{aligned} \Delta y_t &= y_t - y_{t-1} = (1 - B)y_t, \\ \Delta^2 y_t &= \Delta(\Delta y_t) = y_t - 2y_{t-1} + y_{t-2} = (1 - B)^2 y_t, \end{aligned} \quad (1.2.41)$$

.....  
Model ARIMA(p,d,q) lze tedy přepsat do tvaru



$$\varphi(B)(1-B)^d y_t = \theta(B)\varepsilon_t. \quad (1.2.42)$$

Operátor  $\nu(B) = \varphi(B)(1-B)^d$  se nazývá zobecněný autoregresní operátor.

Při výstavbě modelu ARIMA(p,d,q) se postupuje následujícím způsobem. Nejdříve se diferencováním analyzované řady  $y_t$  zkonstruuje stacionární řada  $w_t$ . Pro tuto řadu se vybuduje proces ARMA(p,q). Vlastní konstrukce modelu ARIMA je ještě obvykle zahájena případným transformováním časové řady a pak stanovením řádu diferencování d. Účelem transformace časové řady je linearizovat řadu tak, aby ji bylo možné popsat modelem ARIMA nebo jeho speciálními případy AR, MA nebo ARMA. Linearizací se dosáhne především toho, že náhodné šoky generující řadu mají opravdu charakter bílého šumu s konstantním rozptylem a často navíc s normálním rozdělením. Box a Cox (viz [13]) navrhli transformaci

$$\begin{aligned} z_t &= \frac{y_t^\lambda - 1}{\lambda} && \text{pro } \lambda \neq 0, \\ z_t &= \ln y_t && \text{pro } \lambda = 0 \end{aligned} \quad (1.2.43)$$

kde  $z_t$  označuje transformovanou řadu a  $\lambda$  transformační parametr. Odhad parametru  $\lambda$  je však poměrně komplikovaný (optimalizační procedura je založena na principu tzv. maximální věrohodnosti), proto se často používá transformace (viz např. [54])

$$\begin{aligned} z_t &= y_t^\lambda && \text{pro } \lambda \neq 0 \\ z_t &= \ln y_t && \text{pro } \lambda = 0 \end{aligned} \quad (1.2.44)$$

Prakticky nejvíce používané hodnoty jsou  $\lambda = -1; -0,5; 0; 0,5; 1$ . Jednotlivým hodnotám parametru  $\lambda$  odpovídají po řadě transformace  $\frac{1}{y_t}, \frac{1}{\sqrt{y_t}}, \ln y_t, \sqrt{y_t}, y_t$ .

Přibližné hodnoty parametru  $\lambda$  lze zjistit tak, že časovou řadu rozdělíme na krátké úseky (zhruba 4-12 pozorování) a zjistíme odpovídající průměry  $\bar{y}_j$  a směrodatné odchylky  $s_j$ . Pokud leží body  $[\bar{y}_j, s_j]$  přibližně na horizontální přímce, transformace není

nutná. Leží –li tyto body přibližně na rostoucí přímce, použije se logaritmická transformace.

Po transformaci řady se přejde k určení řádu diferencování. Přehled metod, které se doporučují pro určení řádu diferencování je uveden v [20]. Pro časové řady z praxe se přitom málokdy použije řád diferencování  $d$  větší než 2. Nyní se už sestaví model ARMA(p,q) pro řadu  $w_t = \Delta^d z_t$ .

Nejjednodušším a zároveň prakticky důležitý speciální případ ARIMA(0,1,0) je proces náhodné procházky

$$(1 - B)y_t = \varepsilon_t. \quad (1.2.45)$$

Bez operátoru zpětného posunutí lze zapsat jako

$$y_t = y_{t-1} + \varepsilon_t. \quad (1.2.46)$$

Poznamenejme, že modely ARIMA(p,d,q) mohou kromě trendů vyžadujících stochastické modelování zachytit i čistě deterministické trendy. Model ARIMA(p,d,q) můžeme rozšířit následujícím způsobem

$$\varphi(B)w_t = \delta + \theta(B)\varepsilon_t, \quad (1.2.47)$$

kde  $\delta$  je konstanta. Této definici vyhovují procesy tvaru (plyne z poznatku o diferencování polynomů)

$$\beta_0 + \beta_1 t + \dots + \beta_d t^d + y_t, \quad (1.2.48)$$

kde

$\beta_0 + \beta_1 t + \dots + \beta_d t^d$  je polynomický trend řádu  $d$  a  $y_t$  je proces ARIMA(p,d,q) tvaru  $\varphi(B)w_t = \theta(B)\varepsilon_t$ .

V případě tohoto modelu  $w_t$  představuje proces ARMA(p,q) (1) s nenulovou střední hodnotou.

### 1.2.2.2. Modely časových řad s dlouhou pamětí (ARFIMA(p,d,q))

U stacionárních procesů ARMA klesá ACF exponenciálně s rostoucím zpožděním. To znamená, že náhodné veličiny, které jsou od sebe časově vzdálené, můžeme pokládat za

téměř nekorelované. V praxi se můžeme setkat s časovými řadami, tvořenými stacionárními procesy, jejichž i velmi časově vzdálené náhodné veličiny jsou poměrně silně korelované. Právě v oblasti finančních časových řad se často vyskytují případy, kdy pozorujeme závislost hodnot i v případě velkých časových zpoždění. Časové řady s touto vlastností se označují jako řady s dlouhou pamětí a jejich generující stochastické procesy jako procesy s dlouhou pamětí. Jejich charakteristickou vlastností je, že hodnoty ACF neklesají s rostoucím zpožděním exponenciálně, ale hyperbolicky.

Toto chování je možné modelovat pomocí procesů ARFIMA(p,d,q), které jsou zobecněním modelů ARIMA(p,d,q). V procesech ARFIMA není kladeno žádné omezení na řád diferencování a  $d$  může být libovolné reálné číslo, které se označuje jako frakcionální parametr. Uvažujme nejprve proces ARFIMA(0,d,0).

### Proces ARFIMA(0,d,0)

Proces ARFIMA(0,d,0) je definován jako

$$\Delta^d y_t = (1 - B)^d y_t = \varepsilon_t. \quad (1.2.49)$$

Pro  $d = 0$  je  $y_t = \varepsilon_t$ , tj. bílý šum. Pro  $d = 1$  je  $\Delta y_t = \varepsilon_t$ , tj. náhodná procházka. Je-li  $d$  reálné číslo, platí binomický rozvoj

$$\Delta^d = (1 - B)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-B)^k = 1 - dB - \frac{1}{2}d(1-d)B^2 - \frac{1}{6}d(1-d)(2-d)B^3 - \dots \quad (1.2.50)$$

Základní vlastnosti procesu ARFIMA(0,d,0) lze shrnout následujícím způsobem ([49]):

- Je-li  $d < 0,5$ , pak  $y_t$  je stacionární proces MA nekonečného řádu

$$y_t = \sum_{j=0}^{\infty} \theta_j \varepsilon_{t-j}, \text{ kde } \theta_j = \frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(d)}, \lim_{j \rightarrow \infty} \theta_j = \frac{j^{d-1}}{\Gamma(d)}. \quad (1.2.51)$$

$\Gamma$  je gama funkce (viz např. [27]).

- Je-li  $d > -0,5$ , pak  $y_t$  je invertibilní AR proces nekonečného řádu

$$y_t = \sum_{j=0}^{\infty} \varphi_j y_{t-j} + \varepsilon_t, \text{ kde } \varphi_j = \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(d)}, \lim_{j \rightarrow \infty} \varphi_j = \frac{j^{-d-1}}{\Gamma(-d)}. \quad (1.2.52)$$

- Autokovarianční funkce je dána vztahem

$$\gamma_k = \frac{(-1)^k \Gamma(1-2d)}{\Gamma(k-d+1)\Gamma(1-k-d)}. \quad (1.2.53)$$

- Autokorelační funkce je dána vztahem

$$\rho_k = \frac{\Gamma(1-d)\Gamma(k+d)}{\Gamma(d)\Gamma(k-d+1)}, \rho_1 = \frac{d}{1-d}, \lim_{k \rightarrow \infty} \rho_k = \frac{\Gamma(1-d)}{\Gamma(d)} k^{2d-1}. \quad (1.2.54)$$

- Parciální autokorelační funkce je dána vztahem

$$\rho_{kk} = \frac{d}{k-d}. \quad (1.2.55)$$

Nejdůležitější vlastností z hlediska modelování finančních časových řad je charakter poklesu ACF. Pro  $0 < d < 0,5$  klesá ACF hyperbolicky a platí, že všechny autokorelace jsou kladné a

$$\lim_{n \rightarrow \infty} \sum_{k=-n}^n |\rho_k| = \infty. \quad (1.2.56)$$

V tomto případě se označuje proces jako perzistentní proces nebo také proces s dlouhou pamětí. Rovněž PACF klesá hyperbolicky jako  $\frac{1}{k}$ . Pro  $-0,5 < d < 0$  jsou všechny hodnoty ACF i PACF záporné a platí

$$\lim_{n \rightarrow \infty} \sum_{k=-n}^n |\rho_k| = A \in R. \quad (1.2.57)$$

V tomto případě se proces označuje jako antiperzistentní, nebo také proces se střední pamětí.

### Proces ARFIMA(p,d,q)

Zobecněním modelu ARIMA(p,d,q) je model ARFIMA(p,d,q). Vhodná volba diference  $d$  ( $d$  je libovolné reálné číslo) určí dlouhou paměť procesu a tím i korelační strukturu pro velká zpoždění, zatímco AR a MA parametry určí korelační strukturu pro malá zpoždění.

ARFIMA(p,d,q) proces lze vyjádřit ve tvaru

$$\varphi(B)(1-B)^d y_t = \theta(B)\varepsilon_t \quad (1.2.58)$$

a nazývá se autoregresní frakcionálně integrovaný proces klouzavých průměrů řádu  $p, d, q$ . Vlastnosti procesu ARFIMA( $p,d,q$ ) jsou obdobné jako vlastnosti procesu ARFIMA( $0,d,0$ ).

Při výstavbě modelu ARFIMA je problematika odhadů parametrů komplikovanější než u modelů ARIMA. Frakcionální parametr  $d$  má ve srovnání s ostatními parametry specifický význam, neboť podle jeho hodnoty lze určit, zda se jedná o proces s dlouhou či krátkou pamětí a zda se jedná o stacionární či nestacionární proces. Existují dvě skupiny metod pro získání odhadu tohoto parametru. Na jedné straně jsou to metody, které jej odhadují individuálně, a na druhé straně metody, které jej odhadují současně s ostatními parametry modelu.

Mezi metody, které odhadují frakcionální parametr individuálně, patří např. Hurstův koeficient ([50]) nebo semiparametrický odhad  $d$  ([32]). Pokud časová řada je generována nikoliv pouze frakcionálně integrovaným procesem, ale i procesem, který obsahuje také složku krátké paměti, nastávají určité problémy. Problémy nastávají i v případě krátkých časových řad a při jejich heteroskedasticitě. Také při použití semiparametrického odhadu parametru  $d$  mohou nastat problémy, odhad může být totiž značně vychýlený. V současnosti se proto někteří autoři spíše přiklánějí ke komplexnějšímu přístupu, kdy jsou všechny parametry modelu ARFIMA odhadnuty současně. K odhadu parametrů modelu ARFIMA( $p,d,q$ ) ve tvaru

$$\varphi(B)(1-B)^d y_t = \theta(B)\varepsilon_t \quad (1.2.59)$$

se používá metoda přesné maximální věrohodnosti ([72], [73]). Touto metodou se získají maximálně věrohodné odhady parametrů  $d, \varphi_j, \theta_j$ . Odhady všech parametrů modelu

ARFIMA se dají získat i na základě nelineární metody nejmenších čtverců (viz např. [11]). Odhady získané touto metodou jsou asymptoticky vydatné a normální.

## 1.3. Modely s přenosovou funkcí

Tyto modely pokrývají poměrně širokou oblast časových řad. Zabývají se závislostí mezi časovými řadami a tvorbou modelů, které tyto závislosti popisují. Výstavba modelů s přenosovou funkcí se opírá o Box-Jenkinsovu metodologii výstavby ARIMA modelů, která byla podrobně popsána v kapitole 1.2, tudíž při výstavbě modelu opět postupujeme systematicky: identifikace modelu, odhad parametrů modelu a ověření adekvátnosti modelu.

### 1.3.1. Základní pojmy

#### Dynamický regresní model

Dynamický regresní model je takový model pro časové řady, ve kterém je výstupní řada  $Y_t$  lineárně závislá na současné hodnotě a případně i na minulých hodnotách vstupní řady, resp. více vstupních řad, tedy na hodnotách  $X_t, X_{t-1}, X_{t-2}, \dots$ . Pokud máme více vstupních řad, předpokládáme, že jsou pozorovány ve stejných časových bodech. Výstupní řada je tedy ovlivněna jak současnou tak minulými hodnotami řady vstupní. Avšak vstupní řada není ovlivněna minulými hodnotami řady výstupní. Jsou-li výše uvažované předpoklady splněny, hovoříme o tzv. *jednorovnicovém modelu*.

#### Regresní analýza

Většinou budeme pracovat s klasickým regresním modelem, přičemž budeme používat základní odvození, výsledky a postupy běžné v regresní analýze a budeme se odkazovat na literaturu, kde jsou tyto informace obsaženy (viz např. [4]).

Klasický regresní model, který budeme uvažovat, bude mít většinou tvar

$$Y_t = c + v_0 X_t + N_t, \quad (1.3.1)$$

kde  $c$  je neznámá konstanta,  $v_0$  je neznámý parametr a  $N_t$  je tzv. poruchová řada. Model (1.3.1) je typický jednorovnicový dynamický regresní model, kde  $Y_t$  je výstup a  $X_t$  je

vstup. V klasickém lineárním regresním modelu se předpokládá, že řada  $N_t$  je *bílým šumem*, tedy  $N_t = \varepsilon_t$  a že tyto hodnoty jsou lineárně nezávislé na řadě  $X_t$ .

### **Časově zpožděný model**

Model (1.3.1) popisuje závislost jedné časové řady na druhé ve stejném čase. Často se stává, že výstupní řada závisí nejen na současné hodnotě vstupní řady, ale i na jejích předchozích hodnotách. V tomto případě musíme model (1.3.1) přepsat s respektováním tohoto časového zpoždění. Pokud výstup bude záviset na současné hodnotě a  $n$  minulých hodnotách vstupu, potom příslušný model má tvar

$$Y_t = c + v_0 X_t + v_1 X_{t-1} + v_2 X_{t-2} + \dots + v_n X_{t-n} + N_t \quad (1.3.2)$$

V případě, že bychom nerespektovali časové zpoždění závislosti a použili bychom místo modelu (1.3.2) model (1.3.1), docházelo by ke ztrátě informace, neboť při odhadu  $Y_t$  bychom neuvažovali vliv veličin  $X_{t-1}, X_{t-2}, \dots, X_{t-n}$ . Důsledkem toho by zřejmě odhad reziduálního rozptylu (rozptylu řady  $N_t$ ) byl větší než je nutné a zkonstruované případné předpovědi by nedosahovaly takové přesnosti, jako při respektování časového zpoždění závislosti. Dále by to vedlo zřejmě k vychýlení odhadu parametru  $v_0$  a jeho nekonzistenci.

### **Poruchová řada**

Řadu  $N_t$  v modelu (1.3.1) jsme nazvali poruchovou řadou. Velice často bývají hodnoty této řady navzájem korelovány (hovoříme tedy o autokorelaci). Autokorelace sebou přináší problémy při práci s klasickým regresním modelem. Autokorelaci prvního řádu můžeme pro poruchovou řadu zapsat ve tvaru

$$N_t = \varphi_1 N_{t-1} + \varepsilon_t, \quad (1.3.3)$$

kde  $\varphi_1$  je autoregresní parametr a  $\varepsilon_t$  je bílý šum. Modelování poruchové řady je možné pomocí Box – Jenkinsovy metodologie, kterou jsme stručně popsali v kapitole 1.2.

### 1.3.2. Základní charakteristika modelů s přenosovou funkcí

Modely s přenosovou funkcí spadají do skupiny dynamicky regresních modelů. Tyto modely se nazývají dynamickými ze dvou důvodů. První důvod je ten, že hodnoty výstupní řady  $Y_t$  závisejí na hodnotách poruchové řady, kterou lze popsat ARIMA modelem a druhý důvod je ten, že hodnoty  $Y_t$  mohou záviset jak na současné tak i na minulých hodnotách vstupní řady  $X_t$ . Pokud model obsahuje časově zpožděnou závislost, nazýváme je modely se zpožděním (distributed lag models). V této práci budeme využívat pouze racionální modely se zpožděním (viz [47]).

#### 1.3.2.1. Přenosová funkce

Uvažujme závislost výstupní řady  $Y_t$  na hodnotách vstupní řady  $X_t$ . Tuto závislost můžeme zapsat ve tvaru

$$Y_t = f(X_t), \quad (1.3.4)$$

kde symbol  $f(\cdot)$  značí nějakou funkci. Funkci  $f$  lze označit jako přenosovou funkci – funkce  $f$  přenáší změny v hodnotách řady  $X_t$  na hodnoty řady  $Y_t$ . Soustředíme se pouze na funkce, kde  $f(\cdot)$  bude tzv. lineárně zpožděná funkce (linear distributed lag). Budeme tedy předpokládat, že výstupní řada  $Y_t$  je lineární kombinací současné hodnoty vstupní řady  $X_t$  a jejích minulých hodnot  $X_{t-1}, X_{t-2}, \dots$ . Cílem je najít takový tvar funkce  $f(\cdot)$ , který bude vhodně popisovat chování reálných dat.

Situace nebude ovšem tak jednoduchá jako v modelu (1.3.4). Budeme pracovat s modely typu

$$Y_t = c + f(X_t) + N_t, \quad (1.3.5)$$

Jak již bylo výše řečeno,  $N_t$  představuje poruchovou řadu, která bývá většinou autokorelována. Na řadu  $N_t$  můžeme pohlížet jako na tu část modelu (1.3.5), která vysvětluje tu část variability řady  $Y_t$ , která zůstala nevysvětlena chováním  $X_t$ . V modelu



s přenosovou funkcí se musí odhadnout nejen funkce  $f$ , ale musí se zkonstruovat i model ARIMA pro řadu  $N_t$ .

Nejprve se soustředíme na funkci  $f$ . Pokud budeme uvažovat jednoduchý model (1.3.4) a funkci  $f$ , která je lineární, obdržíme model ve tvaru

$$Y_t = f(X_t) = v_0 X_t + v_1 X_{t-1} + v_2 X_{t-2} + \dots, \quad (1.3.6)$$

kde jednotlivé parametry  $v_j$  se nazývají impulsní váhy (impulse response weights). Ve skutečnosti však tvar přenosové funkce nemusí být lineární, ale předpoklad linearity má však několik výhod, tudíž nadále se tohoto tvaru budeme držet.

Pokud změna v řadě  $X_t$  v časovém okamžiku  $t$  nemá okamžitou odezvu v řadě  $Y_t$ , ale mezi touto změnou a odezvou uplyne několik dalších časových jednotek, hovoříme o tzv. časovém posunu. V tomto případě bude několik prvních parametrů  $v_k$  v modelu (1.3.6) nulových. Časový posun budeme značit písmenem  $b$ .

Použijeme-li zpětného operátoru  $B$  k zapsání modelu (1.3.6), můžeme tento model přepsat do tvaru

$$Y_t = v(B)X_t, \quad (1.3.7)$$

kde  $v(B) = v_0 + v_1 B + v_2 B^2 + v_3 B^3 + \dots$

Jednotlivé parametry  $v_j$  jako funkce  $j$  jsou nazývány impulsní funkce (impulse response function).

### 1.3.2.2. Koyckův model a množina racionálních modelů

Impulsní funkce může mít různý tvar. Některé typy této funkce se však v praxi vyskytují poměrně často. Jeden z nejpoužívanějších modelů je Koyckův model (viz např. [20]).

Uvažujme exponenciální impulsní váhy. Necht'  $\delta_1$  je konstanta, pro kterou platí  $0 < |\delta_1| < 1$ . Předpokládejme nulový časový posun  $b = 0$  a dále předpokládejme, že pro impulsní váhy  $v_j$  platí vztahy:

$$\begin{aligned}
v_1 &= \delta_1 v_0 \\
v_2 &= \delta_1 v_1 \\
v_3 &= \delta_1 v_2 \\
&\cdot \\
&\cdot \\
&\cdot \\
v_k &= \delta_1 v_{k-1} \quad k = 1, 2, 3, \dots
\end{aligned} \tag{1.3.8}$$

Vidíme, že hodnota  $v_0$  slouží jako výchozí hodnota impulsních vah a při každém posunu o jednu časovou jednotku zpět se hodnota vah mění přímo úměrně s násobkem  $\delta_1$ . To znamená, že váhy klesají exponenciálně směrem do minulosti.

Impulsní váhy můžeme rekurentně vyjádřit ve tvaru

$$v_k = \delta_1^k v_0, \quad k \geq 0. \tag{1.3.9}$$

Hodnota  $v_0$  je tedy výchozí hodnotou a hodnota  $\delta_1$  určuje rychlost poklesu.

Vyjádříme-li model (1.3.6) s užitím těchto impulsních vah obdržíme tzv. Koyckův model

$$Y_t = v_0 X_t + \delta_1 v_0 X_{t-1} + \delta_1^2 v_0 X_{t-2} + \dots \tag{1.3.10}$$

Pro čas  $t-1$  obdržíme

$$Y_{t-1} = v_0 X_{t-1} + \delta_1 v_0 X_{t-2} + \delta_1^2 v_0 X_{t-3} + \dots \tag{1.3.11}$$

Po vynásobení rovnice (1.3.11) hodnotou  $\delta_1$  dostaneme

$$\delta_1 Y_{t-1} = \delta_1 v_0 X_{t-1} + \delta_1^2 v_0 X_{t-2} + \delta_1^3 v_0 X_{t-3} + \dots \tag{1.3.12}$$

a po odečtení rovnice (1.3.10) a (1.3.12) máme

$$\begin{aligned}
Y_t - \delta_1 Y_{t-1} &= v_0 X_t, \\
\text{tj. } Y_t &= v_0 X_t + \delta_1 Y_{t-1}.
\end{aligned} \tag{1.3.13}$$

Dostáváme tedy model pouze se dvěma parametry, které je třeba odhadnout. Jakmile odhadneme tyto parametry  $v_0$  a  $\delta_I$ , můžeme spočítat další odhady impulsních vah s využitím rekurentního vztahu (1.3.9).

Pokud budeme uvažovat Koyckův model s nenulovým časovým posunem  $b=d$ , analogickými úpravami jako v případě bez posunu obdržíme model

$$Y_t = v_d X_{t-d} + \delta_I Y_{t-1}. \quad (1.3.14)$$

Vrátíme-li se ke Koyckovu modelu ve tvaru (1.3.13), můžeme pomocí operátoru zpětného posunutí zapsat ve tvaru

$$Y_t = \frac{v_0}{1 - \delta_I B} X_t, \quad (1.3.15)$$

tudíž

$$v(B) = \frac{v_0}{1 - \delta_I B}. \quad (1.3.16)$$

V tomto případě hovoříme o tzv. modelu v racionálním tvaru nebo zkráceně o racionálním modelu.

**Obecný racionální model** můžeme vyjádřit ve tvaru

$$v(B) = \frac{\omega_h(B)B^b}{\delta_r(B)}, \quad (1.3.17)$$

kde

$$\omega_h(B) = \omega_0 + \omega_1 B + \dots + \omega_h B^h \quad (1.3.18)$$

a

$$\delta_r(B) = 1 - \delta_1 B - \dots - \delta_r B^r \quad (1.3.19)$$

a parametr  $b$  představuje časový posun.

Koyckův model (1.3.13) je tedy speciálním případem racionálního modelu, položíme-li  $h = 0, r = 1, \omega_0 = v_0$ .

V případě modelu (1.3.17) výpočet impulsních vah se bude řídit následujícím. Vzorec (1.3.17) přepíšeme do tvaru

$$\delta_r(B)v(B) = \omega_h(B)B^b, \quad (1.3.20)$$

což po rozepsání zapíšeme do tvaru

$$(1 - \delta_1 B - \dots - \delta_r B^r)(v_0 + v_1 B + v_2 B^2 + \dots) = (\omega_0 + \omega_1 B + \dots + \omega_h B^h)B^b. \quad (1.3.21)$$

Pohlížíme-li na (1.3.21) jako na polynom s proměnnou  $B$  a porovnáme koeficienty na obou stranách rovnice u stejných mocnin  $B$ , dojdeme k vyjádření

$$\begin{aligned} v_j &= 0 & j < b \\ v_j &= \delta_1 v_{j-1} + \delta_2 v_{j-2} + \dots + \delta_r v_{j-r} + \omega_0 & j = b \\ v_j &= \delta_1 v_{j-1} + \delta_2 v_{j-2} + \dots + \delta_r v_{j-r} + \omega_{j-b} & j = b+1, b+2, \dots, b+h, \\ v_j &= \delta_1 v_{j-1} + \delta_2 v_{j-2} + \dots + \delta_r v_{j-r} & j > b+h \end{aligned} \quad (1.3.22)$$

což se dá rekurentně zapsat ve tvaru

$$\begin{aligned} v_j &= 0 & j < b \\ v_j &= \sum_{i=1}^r \delta_i v_{j-i} + \omega_{j-b} & j \geq b \end{aligned} \quad (1.3.23)$$

kde  $\omega_{j-b} = 0$  pro  $j-b > h$ .

Pokud známe hodnoty parametrů  $\omega_k$  a  $\delta_k$ , můžeme dle výše uvedeného počítat hodnoty vah  $v_k$ .

### Stabilita modelu

Na hodnotách impulsních vah závisí též stabilita modelu s přenosovou funkcí. Model s přenosovou funkcí je **stabilní**, jestliže řada impulsních vah je absolutně sčitatelná, tzn.  $\sum |v_j| < \infty$ . Tato podmínka platí, pokud řada  $v(B) = v_0 + v_1 B + v_2 B^2 + \dots$  konverguje pro  $|B| \leq 1$ , kde na  $B$  pohlížíme opět jako na algebraickou proměnnou. Vzhledem k tomu, že

$$v(B) = \frac{\omega_h(B)B^b}{\delta_r(B)},$$

podmínka stability vyžaduje, aby výraz  $\frac{\omega_h(B)B^b}{\delta_r(B)}$  byl konečný pro

$|B| \leq 1$ . To znamená, že koeficienty  $\delta$  ve jmenovateli tohoto výrazu musí vyhovovat určitým stabilním podmínkám.

Těmito podmínkami se budeme nyní zabývat. Nejprve zapíšeme charakteristickou rovnici  $\delta(B) = 0$ , kde na  $B$  pohlížíme jako na algebraickou proměnnou. Podmínky stability vyžadují, aby všechny kořeny charakteristické rovnice ležely vně jednotkového kruhu v komplexní rovině. Např. pro  $r=1$  (Koyckův model s parametrem  $r=1$ ), obdržíme podmínku  $|\delta_1| < 1$ . Pro  $r=2$  mají podmínky stability tvar

$$\begin{aligned} |\delta_2| &< 1 \\ \delta_2 + \delta_1 &< 1 \\ \delta_2 - \delta_1 &< 1 \end{aligned} \quad (1.3.24)$$

Modely s  $r > 2$  se většinou v praxi nepoužívají. Hledání kořenů charakteristické rovnice je v tomto případě již daleko složitější a využívá se proto při něm většinou matematický software (MAPLE, MATHEMATICA, DERIVE...).

### 1.3.2.3. Nejčastější tvary přenosové funkce

V následující tabulce jsou uvedeny nejobvyklejší tvary přenosové funkce. Uvažujeme časový posun  $b = 0$ , neboť dosažené výsledky lze snadno přepsat pro libovolnou hodnotu  $b$ .

<b>(b,r,h)</b>	<b>Tvar přenosové funkce</b>
(0,0,0)	$v(B)X_t = \omega_0 X_t$
(0,0,1)	$v(B)X_t = (\omega_0 + \omega_1 B)X_t$
(0,0,2)	$v(B)X_t = (\omega_0 + \omega_1 B + \omega_2 B^2)X_t$
(0,1,0)	$v(B)X_t = \frac{\omega_0}{(1 - \delta_1 B)} X_t$
(0,1,1)	$v(B)X_t = \frac{(\omega_0 + \omega_1 B)}{(1 - \delta_1 B)} X_t$
(0,1,2)	$v(B)X_t = \frac{(\omega_0 + \omega_1 B + \omega_2 B^2)}{(1 - \delta_1 B)} X_t$
(0,2,0)	$v(B)X_t = \frac{\omega_0}{(1 - \delta_1 B - \delta_2 B^2)} X_t$
(0,2,1)	$v(B)X_t = \frac{(\omega_0 + \omega_1 B)}{(1 - \delta_1 B - \delta_2 B^2)} X_t$
(0,2,2)	$v(B)X_t = \frac{(\omega_0 + \omega_1 B + \omega_2 B^2)}{(1 - \delta_1 B - \delta_2 B^2)} X_t$

Podívejme se na význam jednotlivých částí přenosové funkce. Výraz v čitateli  $\omega_h(B)$  představuje tu část impulsních vah, která nemá charakteristický průběh - jsou to hodnoty parametrů  $\omega_j$ , jejichž chování lze ztěží nějak teoreticky popsat. Jmenovatel zlomku  $\delta_r(B)$  je ta část impulsních vah, která určuje jejich charakteristický průběh. Je-li  $r=0$  můžeme předpokládat, že impulsní váhy nemají žádný charakteristický průběh. V případě  $r=1$ , můžeme očekávat exponenciální průběh impulsních vah a v případě  $r=2$  klesající váhy buď exponenciálně nebo po sinusoidě s klesající amplitudou - záleží na tom, kde leží kořeny polynomu

$$\delta_2(B) = 1 - \delta_1 B - \delta_2 B^2 = 0. \quad (1.3.25)$$

Pokud jsou kořeny reálné, váhy klesají exponenciálně. Jsou-li kořeny komplexní, váhy klesají po sinusoidě s klesající amplitudou.

#### 1.3.2.4. Vícenásobný vstup

Doposud jsme v racionálním modelu uvažovali pouze jednu vstupní řadu  $X_t$ . Nyní uvažovaný model rozšíříme o další vstupní řady a tak získáme vícerozměrný racionální model. Předpokládejme, že máme  $M$  vstupních řad  $X_{1,t}, X_{2,t}, \dots, X_{M,t}$ , potom racionální model bude mít následující tvar

$$Y_t = \sum_{i=1}^M v_i(B) X_{i,t} = \sum_{i=1}^M \frac{\omega_{i,h_i}(B) B^{b_i}}{\delta_{i,r_i}(B)} X_{i,t}, \quad (1.3.26)$$

kde

$$v_i(B) = v_{i,0} + v_{i,1}B + v_{i,2}B^2 + \dots = \frac{\omega_i(B) B^{b_i}}{\delta_i(B)}, \quad i = 1, 2, \dots, M$$

$b_i$  je časový posun pro vstup  $X_{i,t}$ ,  $i = 1, 2, \dots, M$

$$\omega_i(B) = \omega_{i,0} + \omega_{i,1}B + \dots + \omega_{i,h_i} B^{h_i}, \quad i = 1, 2, \dots, M$$

$$\delta_i(B) = 1 - \delta_{i,1}B - \delta_{i,2}B^2 - \dots - \delta_{i,r_i} B^{r_i}, \quad i = 1, 2, \dots, M$$

$h_i$  je řád  $\omega_i(B)$  pro  $i = 1, 2, \dots, M$

$r_i$  je řád  $\delta_i(B)$  pro  $i = 1, 2, \dots, M$ .

Rovnice (1.3.26) ukazuje racionální model s  $M$  vstupy a  $M$  – rozměrnou přenosovou funkcí. Zapišeme nyní celkový tvar racionálního modelu, včetně konstanty  $c$  a poruchové složky  $N_t$  jako

$$Y_t = c + \sum_{i=1}^M \frac{\omega_{i,h_i}(B)B^{b_i}}{\delta_{i,r_i}(B)} X_{i,t} + N_t. \quad (1.3.27)$$

Poruchová složka se řídí obecným ARIMA modelem. ARIMA modely jsme popsali v kapitole 1.2. S ohledem na teorii ARIMA modelů můžeme psát

$$\phi(B)\Phi(B^L)\Delta^d \Delta_L^D N_t = \theta(B)\Theta(B^L)\varepsilon_t. \quad (1.3.28)$$

Rovnici (1.3.28) převedeme na tvar

$$N_t = \frac{\theta(B)\Theta(B^L)}{\phi(B)\Phi(B^L)\Delta^d \Delta_L^D} \varepsilon_t. \quad (1.3.29)$$

Pokud za  $N_t$  dosadíme do rovnice (1.3.27), získáme celkový zápis dynamického regresního modelu

$$Y_t = c + \sum_{i=1}^M \frac{\omega_{i,h_i}(B)B^{b_i}}{\delta_{i,r_i}(B)} X_{i,t} + \frac{\theta(B)\Theta(B^L)}{\phi(B)\Phi(B^L)\Delta^d \Delta_L^D} \varepsilon_t. \quad (1.3.30)$$

Podívejme se nyní na několik konkrétních tvarů modelu (1.3.30), které se v praxi často vyskytují ([66]).

1. Vícenásobný regresní model se dvěma vysvětlujícími časovými proměnnými

$$Y_t = c + \omega_{1,0}X_{1,t} + \omega_{2,0}X_{2,t} + \varepsilon_t. \quad (1.3.31)$$

Jedná se zjevně o speciální případ dynamického regresního modelu, ve kterém,  $M = 2$ ,  $b_1 = b_2 = h_1 = h_2 = r_1 = r_2 = 0$  a  $N_t = \varepsilon_t$  je bílý šum.

2. Regresní model s jednou vysvětlující časovou proměnnou, ve kterém  $Y_t$  závisí na této proměnné ve stejném čase  $t$  a dále na jejích hodnotách v čase  $t-1$ ,  $t-2$ , ... Jedná se tedy o model s časovým zpožděním. Tento typ modelu se často používá v ekonometrii.

$$Y_t = c + \omega_0 X_t + \omega_1 X_{t-1} + \omega_2 X_{t-2} + \dots + \varepsilon_t. \quad (1.3.32)$$

Jde o případ dynamického regresního modelu, ve kterém  $M = 1, b = r = 0$ ,  $h$  je počet časově zpožděných hodnot vstupní řady  $X_t$  a  $N_t = \varepsilon_t$  je opět bílý šum.

3. Regresní model s autokorelovanou poruchovou složkou. Uvažujme pro poruchovou složku  $N_t$  např. model AR(1). Potom

$$(1 - \phi_1 B)N_t = \varepsilon_t \quad (1.3.33)$$

a model

$$Y_t = c + \omega_0 X_t + \omega_1 X_{t-1} + \omega_2 X_{t-2} + \dots + N_t \quad (1.3.34)$$

můžeme přepsat do tvaru

$$Y_t = c + \omega_0 X_t + \omega_1 X_{t-1} + \omega_2 X_{t-2} + \dots + \frac{1}{1 - \phi_1 B} \varepsilon_t. \quad (1.3.35)$$

Jedná se o případ dynamického regresního modelu, ve kterém  $M = 1, b = r = 0$ ,  $h$  je počet časově zpožděných hodnot vstupní řady  $X_t$ ,  $\phi(B) = 1 - \phi_1 B, \theta(B) = 1, \Phi(B^L) = 1, D = 0, d = 0$  a  $\varepsilon_t$  je bílý šum.

### 1.3.3. Identifikace modelu

V této kapitole jsou popsány základní kroky při identifikaci modelu s přenosovou funkcí. Jsou zde uvedeny dvě ze čtyř základních metod, které se pro tento účel používají. Jedná se o identifikační metodu pomocí *lineární přenosové funkce (linear transfer function –LTF*, viz [68]) a dále o metodu založenou na studiu *vzájemné korelační funkce (cross correlation function)* a na tzv. *zbělení* vstupní a výstupní řady (viz [14]).

Uvažujme dynamický regresní model ve tvaru

$$Y_t = c + v(B)X_t + N_t \quad (1.3.36)$$

kde  $X_t$  je vstupní řada,  $Y_t$  je výstupní řada,  $c$  je konstanta,  $v(B)X_t$  je přenosová funkce a  $N_t$  je poruchová složka, která se obecně řídí ARIMA procesem a je nezávislá na vstupu  $X_t$ . Pro tento model se snažíme nejdříve najít co nejjednodušší vyjádření polynomu



$v(B)$  ve tvaru podílu dvou dalších polynomů  $\frac{\omega_h(B)B^b}{\delta_r(B)}$  co nejnížšího stupně. Druhým krokem při výstavbě tohoto modelu je nalezení vhodného ARIMA procesu, který by co nejlépe popisoval chování poruchové řady  $N_t$ . Obecně se uvažuje libovolný ARIMA model, tedy pro  $N_t$  budeme předpokládat  $N_t = ((\theta(B)\Theta(B^L)/\varphi(B)\Phi(B^L)\Delta^d\Delta_L^D)) \varepsilon_t$ , kde  $\varepsilon_t$  je normálně rozdělený bílý šum. Po těchto dvou krocích následuje odhad parametrů (kapitola 1.3.4), fáze ověřování adekvátnosti modelu (kapitola 1.3.5) a předpovědi v modelu (kapitola 1.3.6).

### 1.3.3.1. Předběžné kroky

Než začneme s vlastní identifikací modelu, je třeba provést určité předběžné kroky, které nám pomohou při vlastní identifikaci.

1. *Kontrola dat, odlehlá pozorování.* Nejprve je vhodné si prohlédnout průběh uvažované řady na obrázku. Z průběhu řady je většinou možné určit, zda řada obsahuje trend a sezónnost a dále to, zda řada je výrazně nestacionární. Zobrazení-li průběh obou řad (vstupu i výstupu) do jednoho grafu, je často vidět jejich případná závislost a možný časový posun mezi oběma řadami. Graf nám může napovědět, zda časová řada neobsahuje odlehlá pozorování, intervence či pouze některé chybné údaje, které mohou vzniknout např. při opisování dat, jejich agregací apod. Zjistíme-li přítomnost odlehlých pozorování, musíme s nimi počítat v celém procesu výstavby modelu. Podrobnější informace o intervencích a odlehlých pozorováních můžeme najít v pracích L.Marka ([62], [63], [64], [66]).
2. *Stacionarita.* Jedním z důležitých předpokladů při výstavbě racionálních modelů s přenosovou funkcí je předpoklad stacionarity vstupní a výstupní řady. Tento předpoklad nebývá v praxi často splněn a proto je třeba před vlastním procesem identifikace modelu převést tyto časové řady na stacionární. K stabilizaci rozptylu se většinou používá Box-Coxova transformace, kterou jsme uvedli v kapitole 1.2. Pro dosažení stacionarity se většinou používají běžné, případně sezónní diference.

3. *ARIMA modely*. Před vlastní konstrukcí modelu s přenosovou funkcí bývá vhodné sestrojít ARIMA model jak pro vstupní tak pro výstupní řadu. Důvody jsou následující:

- Sestrojíme-li ARIMA model pro výstupní řadu a porovnáme-li ho s později zkonstruovaným modelem s přenosovou funkcí, může se stát, že pomocí ARIMA modelu dosáhneme lepších výsledků – rezidua mají menší rozptyl a i zkonstruované předpovědi jsou přesnější. Musíme ovšem brát na zřetel i fakt, že je mnohdy užitečné modelovat závislost výstupní řady na řadě vstupní a pak samozřejmě můžeme přistoupit i na složitější či nepatrně „horší“ model s přenosovou funkcí.
- Konstrukce předpovědi pro vstupní řadu. Chceme-li totiž sestrojít v modelu s přenosovou funkcí předpovědi pro výstupní řadu, potřebujeme při jejich výpočtu předpovědi pro řadu vstupní.
- Tvorba ARIMA modelu může poukázat na některé další informace o povaze dat, které pak využijeme při tvorbě modelu s přenosovou funkcí.
- Výsledky dosažené při konstrukci ARIMA modelu pro vstupní řadu můžeme později v modelu s přenosovou funkcí využít při konstrukci ověřovacích kritérií, výpočet některých charakteristik (směrodatná odchylka předpovědí, intervaly spolehlivosti).

4. *Zpětná vazba*. Jak již bylo řečeno výstupní řada  $Y_t$  je lineárně závislá na současné hodnotě a případně i na minulých hodnotách vstupní řady. Důležitým předpokladem je směr této uvažované závislosti, protože vstupní řada naopak nesmí být ovlivněna minulými hodnotami řady výstupní. Pokud existuje závislost vstupní řady na minulých hodnotách řady výstupní, hovoříme o tzv. *zpětné vazbě (feedback)*. Na řady, které vykazují efekt zpětné vazby není možné aplikovat model s přenosovou funkcí, neboť by celý proces výstavby modelu vedl k nekonzistentním odhadům parametrů. Přítomnost zpětné vazby mezi řadami je třeba testovat. Příslušné testy jsou obsaženy v některých softwarech jako např. v SCA nebo v SASu. Problematikou zpětné vazby se zabývá např. publikace [87].

### 1.3.3.2. Identifikace modelu metodou LTF

Zkratka LTF značí „linear transfer function“, tedy lineární přenosovou funkci. Metodu LTF zobecnil Alan Pankratz ([68]) a modely s přenosovou funkcí jsou v ní označovány jako dynamické regresní modely. Od této chvíle tedy nebudeme rozlišovat pojmy „model s přenosovou funkcí“ a „dynamický regresní model“. Celá metoda se dá shrnout do následujících kroků:

- a) určení předběžného tvaru přenosové funkce
- b) předběžné určení řádu běžného příp. i sezónního autoregresního modelu pro poruchovou řadu  $N_t$
- c) odhady koeficientů
- d) ověření stacionarity poruchové řady  $N_t$  - je-li stacionarita splněna, pokračujeme v procesu identifikace, není-li splněna, je třeba diferencovat vstupní a případně i výstupní řadu za účelem dosažení stacionarity a poté začít znovu od bodu a)
- e) určení vhodného modelu pro přenosovou funkci
- f) určení vhodného ARMA modelu pro poruchovou řadu
- g) odhady koeficientů
- h) ověření adekvátnosti modelu – je-li vše v pořádku, končíme s výstavbou modelu, případně můžeme pokračovat tvorbou předpovědí; není-li některý z modelů v e) či f) v pořádku, navracíme se do tohoto bodu a celý model předefinujeme.

**Ad a)** V této fázi se pracuje s předběžným modelem. Zkonstruujeme tedy model ve tvaru

$$Y_t = c + v(B)X_t + N_t = c + v_0X_t + v_1X_{t-1} + \dots + v_KX_{t-K} + N_t \quad (1.3.37)$$

kde  $K$  je maximální možný časový posun. Model (1.3.37) se označuje jako lineární přenosová funkce. Ze zápisu (1.3.37) je zřejmé, že slabinou této konstrukce je vhodná volba maximálního časového posunu  $K$ . Při stanovení hodnoty  $K$  sehrává roli i zkušenost statistika a věcná znalost celé problematiky a souvislostí mezi modelovanými řadami. Pro vhodnou volbu konstanty  $K$  existuje ale i metoda založená na vícerozměrných ARMA modelech.

Dále v modelu (1.3.37) musíme odhadnout parametry. Vzhledem k tomu, že se v podstatě jedná o lineární regresní model, je možné tento odhad provést ve většině statistických softwarů, ale nesmíme zapomínat, že se jedná o model s výraznou multikolinearitou. Většinou je třeba provést diference vstupních řad, přičemž dochází ke ztrátě pozorování. Poté se odhadnuté hodnoty vah  $v_k$  porovnají se svými teoretickými protějšky. Porovnání empirických a teoretických hodnot nám umožní stanovit racionální tvar přenosové funkce.

**Ad b)** V tomto kroku se soustředíme na odhalení případné autoregrese v poruchové řadě  $N_t$ . Jedná se v podstatě o aproximaci případné autokorelační struktury poruchové řady  $N_t$ . Uvažujeme pouze autoregresi nízkého řádu, neboť autoregrese vyšších řádů se vyskytuje zřídka. Hlavní důvod tvorby předběžného AR modelu je ten, že pokud bychom zcela ignorovali autokorelaci poruchové řady  $N_t$ , narazíme na problémy při odhadu vah  $v_k$ . Jednak neobdržíme eficientní odhady těchto vah, jednak dílčí regresní t-testy, které používáme pro určení významnosti jednotlivých regresních koeficientů, nebudou mít žádnou vypovídací schopnost. Toto by mělo vliv na rozhodování, které „minulé“ hodnoty  $X_t$  zařadit do modelu. Navíc, budeme-li schopni alespoň zjednodušeně popsat autokorelační strukturu poruchové řady  $N_t$ , umožní nám to lépe pochopit chování této řady a využít těchto předběžných výsledků pro vybudování odpovídajícího ARIMA modelu v další fázi práce.

Pro poruchovou řadu uvažujme nejjednodušší AR model ve tvaru

$$(1 - \phi_1(B))(1 - \Phi_1(B^L))N_t = \varepsilon_t \quad (1.3.38)$$

O hodnotách  $\varepsilon_t$  předpokládáme, že se jedná o bílý šum. Je zřejmé, že předpoklad o bílém šumu nemusí platit. Ve skutečnosti i náhodná složka  $\varepsilon_t$  může vykazovat autokorelaci. Dosadíme-li (1.3.38) do (1.3.37), obdržíme

$$Y_t = c + v_0 X_t + v_1 X_{t-1} + \dots + v_K X_{t-K} + \frac{1}{(1 - \phi_1(B))(1 - \Phi_1(B^L))} \varepsilon_t \quad (1.3.39)$$

Rovnice (1.3.39) nám umožní odhadnout zároveň váhy  $v_k$  i autoregresní parametry  $\phi_1, \phi_1$ . Získáme tím předběžný tvar modelu s přenosovou funkcí.

**Ad c)** Odhadnutí parametrů v modelu (1.3.39) se řeší použitím vhodného softwaru. Software vhodný pro odhad příslušných parametrů je např. SCA nebo SAS. Oba software jsou přímo zaměřeny na práci s časovými řadami a mají modely s přenosovou funkcí přímo zabudovány v nabídce jako procedury. Oba zmíněné programy umožňují zvolit mezi několika metodami výpočtu.

SAS nabízí tři metody. Jde o metodu ML = „maximum likelihood“ (metoda maximální věrohodnosti, při které je věrohodnostní funkce počítána na základě Kalmanových filtrů), ULS = „unconditional least-squares“ (nepodmíněná metoda nejmenších čtverců) a metodu CLS = „conditional least-squares“ (podmíněná metoda nejmenších čtverců)

SCA nabízí dvě metody odhadu. Jedná se o metodu ML založenou na podmíněné věrohodnostní funkci ([14]). Tato metoda dle autorů paketu SCA nevede vždy k dosažení nejlepších odhadů, neboť produkuje vychýlené odhady pro MA model pro poruchovou řadu. Používá se především v případě, že ARIMA model pro poruchovou řadu neobsahuje MA složku. Dále je v SCA nabízena modifikace této procedury ([48]), která je sice výpočetně náročnější, ale produkuje odhady MA koeficientů s lepšími vlastnostmi.

**Ad d)** K ověření stacionarity poruchové řady  $N_t$  se používá běžných postupů známých z teorie výstavby ARIMA modelů. Zjistíme-li nestacionaritu poruchové řady, je třeba provést diferencování jak vstupní řady, tak řady výstupní. Toto diferencování se provádí před dalšími kroky ihned po případném zjištění nestacionarity. Ověřování stacionarity je založeno na studiu průběhu výběrové autokorelační funkce řady  $\hat{N}_t$ . Protože tuto řadu však neznáme, musíme nejprve získat její odhad ve tvaru

$$\hat{N}_t = Y_t - \hat{c} - \hat{v}(B)X_t \quad (1.3.40)$$

Řada  $\hat{N}_t$  je např. v SCA spočítána jako vedlejší produkt při odhadu modelu (1.3.39). Jakmile máme hodnoty řady  $\hat{N}_t$  k dispozici, můžeme studovat chování její výběrové autokorelační funkce a rozhodnout o stacionaritě této řady.

**Ad e)** Stěžejní částí výstavby dynamického regresního modelu je určení tvaru přenosové funkce. Většinou se stačí omezit na modely s parametry  $(b,r,h)$ , které jsou nejvýše rovny 2. Tvary těchto nejčastěji používaných přenosových funkcí jsme uvedli v kapitole 1.3.2.3. Pokud bychom graficky znázornily průběhy těchto teoretických přenosových funkcí a porovnali je vizuálně s našimi výběrovými hodnotami, můžeme na základě tohoto porovnání určit přesný tvar racionální přenosové funkce. Pokud vhodný tvar přenosové funkce není zjevný, používají se tzv. identifikační pravidla.

*Identifikační pravidla.* Pravidla pro identifikaci přenosové funkce, které uvedeme, jsou uvažována pro případ teoretického tvaru této funkce. V praxi však pracujeme s výběrovými hodnotami, které budeme s teoretickými hodnotami porovnávat. Je tedy zřejmé, že zde bude docházet k určitým odchýlkám a může se stát, že v některých případech nebude vyhodnocení zcela jednoznačné. Pravidla pro identifikaci slouží pouze jako pomocný prostředek, což znamená, že námi vybraný tvar přenosové funkce nemusí být vždy v souladu s těmito pravidly (jedná se o podobný problém, jako je volba vysvětlujících proměnných v regresním modelu – občas zařadíme do modelu proměnnou spíše na základě věcných důvodů, než na základě statistických analýz).

Znovu připomeňme, že čítec  $\omega_h(B)B^b$  v racionální přenosové funkci pokrývá tři složky: časový posun  $b$ ; váhy  $v_k$ , které nemají charakteristický průběh (jejich průběh je spíše chaotický), a váhy  $v_k$ , které můžeme označit jako počáteční charakteristické váhy (váhy, od kterých má již přenosová funkce charakteristický průběh). Ve jmenovateli  $\delta_r(B)$  v racionální přenosové funkci jsou váhy  $v_k$ , které mají zcela charakteristický průběh.

Nyní budou následovat **pravidla pro identifikaci přenosové funkce:**

1. Určení časového posunu  $b$ . Tento krok je poměrně jednoduchý, neboť hodnota  $b$  je rovna počtu počátečních vah  $v_k$ , které se rovnají nule (myšleno tak, že tyto váhy jsou statisticky nevýznamné), přičemž sledování provádíme od první váhy  $v_0$ . Hodnotu  $b$  lze většinou určit z grafu.
2. Určení hodnoty  $r$ . Hodnota  $r$  je řádem polynomu  $\delta_r(B)$  ve jmenovateli racionální přenosové funkce. Při určování  $r$  se budeme řídit těmito pravidly:
  - pokud je průběh vah nesystematický a připomíná spíše skupinu jednotlivých hodnot následovanou dalšími váhami blízkými nebo rovnými nule, volíme  $r=0$ ,
  - jestliže váhy klesají exponenciálně, buď okamžitě od první hodnoty nebo k tomuto klesání předchází několik hodnot, které nevykazují charakteristický průběh, volíme  $r=1$ ,
  - jestliže váhy klesají jako směsice exponenciálních křivek či po sinusoidě s klesající amplitudou (buď okamžitě či po několika nenulových hodnotách), volíme  $r=2$ ,
  - modely s hodnotami  $r > 2$  se v praxi vyskytují velice zřídka.
3. Určení počátečních hodnot. Je-li  $r > 0$ , je třeba určit tzv. počáteční hodnoty, po kterých mají již váhy charakteristický průběh. Ačkoliv váhy již od jistého bodu vykazují určité systematické chování, potřebují jednu nebo více „startovacích hodnot“ na „ukotvení“. Počet těchto počátečních hodnot je roven parametru  $r$ . Při určování počátečních hodnot postupujeme tak, že nalezneme v absolutní hodnotě největší váhu, tj.  $\max|v_k|$ , od které začíná charakteristický průběh. Počet počátečních hodnot je pak dán hodnotou  $r$ . Je-li  $r=1$ , je situace jednoduchá, neboť existuje pouze jedna počáteční hodnota a tou je přímo ona nalezená maximální váha. Je-li  $r=2$ , existují dvě počáteční hodnoty a jedná se o nalezenou maximální váhu a o váhu s indexem o jedničku vyšším.
4. Určení hodnoty  $u$ . Tato hodnota je dána počtem nesystematických vah  $v_k$ .

- Je-li  $r > 0$ ,  $u$  je rovno počtu nenulových (signifikantních) vah vlevo od počáteční hodnoty s nejnižším časovým posunem,
  - je-li  $r=0$ , všechny nenulové váhy mají nesystematický průběh.
5. Určení hodnoty  $h$ , kde  $h$  je řád polynomu  $\omega_h(B)$  v čitateli racionální přenosové funkce. Platí  $h = u + r - 1$ . Je třeba si uvědomit, že  $h$  je řád polynomu  $\omega_h(B)$ , nikoliv počet vah v  $\omega_h(B)$ . Celkový počet vah je totiž o jedničku vyšší než  $h$ , protože váhy jsou indexovány od nuly.

### 1.3.3.3. Identifikace na základě vzájemného vztahu CCF a vah přenosové funkce

Pro určení tvaru přenosové funkce lze využít i klasický postup navržený v knize Box a Jenkins ([14]), který je založený na vzájemném vztahu vzájemné korelační funkce a vah v přenosové funkci.

Definujme tedy **vzájemnou korelační funkci**. Vzájemná korelační funkce (dále jen CCF z anglického Cross – Correlation Function) slouží jako míra intenzity lineární závislosti dvou náhodných veličin. Neurčuje však jenom intenzitu této závislosti, ale určuje zároveň i její směr.

Předpokládejme, že  $X_t$  a  $Y_t$  jsou dva stochastické procesy,  $t = 0, \pm 1, \pm 2, \dots$ . Procesy  $X_t$  a  $Y_t$  nazveme *vzájemně stacionární*, jestliže  $X_t$  a  $Y_t$  jsou každý stacionárním procesem a vzájemná kovarianční funkce mezi  $X_t$  a  $Y_s$  ( $\text{cov}(X_t, Y_s)$ ) je pouze funkcí časového rozdílu  $s-t$ . Tedy

$$\gamma_{XY}(k) = E(X_t - \mu_X)(Y_{t+k} - \mu_Y), \quad k = \dots -1, 0, 1, \dots \quad (1.3.41)$$

kde  $\mu_X$  a  $\mu_Y$  jsou střední hodnoty  $X_t$  a  $Y_t$ . Z rovnice (1.3.41) snadno obdržíme vzájemnou korelační funkci jako

$$\rho_{XY}(k) = \frac{\gamma_{XY}(k)}{\sigma_X \sigma_Y}, \quad k = \dots -1, 0, 1, \dots \quad (1.3.42)$$

kde  $\sigma_X$  a  $\sigma_Y$  jsou směrodatné odchylky  $X_t$  a  $Y_t$ . Existuje tedy velmi úzká souvislost mezi vzájemnou kovarianční funkcí a vzájemnou korelační funkcí na jedné straně a



autokovarianční funkcí a autokorelační funkcí (obě tyto funkce byly definovány v kapitole 1.2) na straně druhé. Je zřejmé, že platí  $\gamma_{XX}(k) = \gamma_X(k)$  a  $\rho_{XX}(k) = \rho_X(k)$ . Avšak vzájemná korelační funkce není symetrická, tudíž píšeme  $\rho_{XY}(k) \neq \rho_{XY}(-k)$ . Z rovnice

$$\gamma_{XY}(k) = E(X_t - \mu_X)(Y_{t+k} - \mu_Y) = E(Y_{t+k} - \mu_Y)(X_t - \mu_X) = \gamma_{YX}(-k) \quad (1.3.43)$$

je vidět, že platí  $\rho_{XY}(k) = \rho_{YX}(-k)$ . Potvrzuje se, že CCF měří nejen intenzitu závislosti, ale i její směr. Proto musíme hodnoty  $\rho_{XY}(k)$  počítat vždy jak pro  $k > 0$  tak pro  $k < 0$ .

Stejně jako u autokovarianční a autokorelační funkce nepracujeme ani v tomto případě s teoretickými hodnotami, ale s odhadnutými. Při odhadu vzájemně korelační funkce použijeme jako v kapitole 1.2 výběrové charakteristiky:

$$\hat{\rho}_{XY}(k) = \frac{\hat{\gamma}_{XY}(k)}{S_X S_Y}, \quad k = \dots -1, 0, 1, \dots \quad (1.3.44)$$

kde

$$\begin{aligned} \hat{\gamma}_{XY}(k) &= \sum_{t=1}^{n-k} \frac{(X_t - \bar{X})(Y_{t+k} - \bar{Y})}{n}, \quad k \geq 0 \\ &= \sum_{t=1-k}^n \frac{(X_t - \bar{X})(Y_{t+k} - \bar{Y})}{n}, \quad k < 0 \end{aligned} \quad (1.3.45)$$

a  $\bar{X}$  a  $\bar{Y}$  jsou aritmetické průměry řad  $X_t$  a  $Y_t$  a  $S_X$  a  $S_Y$  jejich výběrové směrodatné odchylky. V praxi potřebujeme zjistit, které koeficienty CCF  $\rho_{XY}(k)$  jsou nulové. Test provedeme podobně jako v kapitole 1.2 v případě autokorelační funkce. Pokud řada  $X_t$  je bílý šum, porovnáme výběrovou vzájemně korelační funkci s její aproximativní směrodatnou odchylkou  $\frac{1}{\sqrt{n-k}}$  a otestujeme tak hypotézu, že řady  $X_t$  a  $Y_t$  jsou nekorelované. Vzhledem k tomu, že řada  $X_t$  nebývá většinou bílým šumem, musíme ji do tohoto tvaru nejprve upravit.

Nyní uvedeme postup při identifikaci modelu s přenosovou funkcí na základě vztahu CCF a vah přenosové funkce.

### Postup při identifikaci modelu s přenosovou funkcí na základě vztahu CCF a vah přenosové funkce:

1. Zbělíme vstupní řadu

$$\phi_X(B)X_t = \theta_X(B)\alpha_t \quad (1.3.46)$$

(z (1.3.46) je zřejmé, že předpokládáme, že vstupní řada se řídí obecným ARMA modelem), tzn. pro řadu  $\alpha_t$  můžeme psát

$$\alpha_t = \frac{\phi_X(B)X_t}{\theta_X(B)} \quad (1.3.47)$$

kde  $\alpha_t$  je bílý šum s nulovou střední hodnotou a rozptylem  $\sigma_\alpha^2$ . Řada  $\alpha_t$  v (1.3.47) se nazývá zbělená vstupní řada (prewhitened input series).

2. Spočítáme filtrovanou výstupní řadu. Výstupní řadu  $Y_t$  musíme transformovat s použitím zběleného modelu na řadu

$$\beta_t = \frac{\phi_X(B)Y_t}{\theta_X(B)} \quad (1.3.48)$$

3. Spočítáme výběrovou vzájemnou korelační funkci  $\hat{\rho}_{\alpha\beta}(k)$  mezi  $\alpha_t$  a  $\beta_k$ , což nám umožní odhadnout  $v_k$

$$\hat{v}_k = \frac{\hat{\sigma}_\beta}{\hat{\sigma}_\alpha} \hat{\rho}_{\alpha\beta}(k) \quad (1.3.49)$$

Významnost hodnot CCF a tím pádem i významnost vah  $\hat{v}_k$  otestujeme porovnáním se směrodatnou odchylkou  $\frac{1}{\sqrt{n-k}}$ .

4. Určíme  $b, \delta_r(B) = 1 - \delta_1 B - \dots - \delta_r B^r, \omega_h(B) = \omega_0 + \omega_1 B + \dots + \omega_h B^h$  porovnáním hodnot  $\hat{v}_k$  s teoretickými hodnotami  $v_k$ . Jakmile známe hodnoty  $b, r$  a  $h$ , můžeme zkonstruovat průběžné odhady  $\hat{\omega}_j$  a  $\hat{\delta}_j$  na základě rovnic (1.3.22). Předběžný odhad přenosové funkce  $v(B)$  má poté tvar

$$\hat{v}(B) = \frac{\hat{\omega}_h(B)}{\hat{\delta}_r(B)} B^b \quad (1.3.50)$$

5. Poté, co získáme předběžný tvar přenosové funkce, můžeme odhadnout poruchovou řadu jako

$$\hat{N}_t = Y_t - \hat{v}(B)X_t = Y_t - \frac{\hat{\omega}_h(B)}{\hat{\delta}_r(B)} B^b X_t \quad (1.3.51)$$

Identifikační proces je stejný jako v případě identifikace obecného ARMA modelu (využíváme především vlastností ACF, PACF, IACF a dalších nástrojů). Obdržíme obecný model

$$\phi(B)N_t = \theta(B)\varepsilon_t \quad (1.3.52)$$

kde  $\varepsilon_t$  je bílý šum.

6. Z rovnice (1.3.50) a (1.3.52) obdržíme tvar modelu s přenosovou funkcí

$$Y_t = \frac{\omega(B)}{\delta(B)} X_{t-b} + \frac{\theta(B)}{\phi(B)} \varepsilon_t \quad (1.3.53)$$

Po celou dobu konstrukce předpokládáme stacionaritu řad  $X_t, Y_t$  a  $N_t$ . Tento předpoklad však nemusí být vždy splněn. V takovém případě je třeba aplikovat vhodnou transformaci k dosažení stacionarity.

### 1.3.4. Odhady parametrů modelu

Po identifikaci modelu následuje další fáze výstavby modelu, a to je odhad parametrů. Vzhledem k tomu, jak je model s přenosovou funkcí definován, je zřejmé, že odhad parametrů modelu můžeme teoreticky rozčlenit minimálně do dvou kroků, a to na odhad parametrů přenosové funkce a na odhad parametrů poruchové řady. V konečné fázi se pochopitelně odhaduje celý model najednou. Budeme předpokládat klasický model s přenosovou funkcí danou trojicí parametrů  $(b, r, h)$  a poruchovou složkou popsanou ARIMA modelem řádu  $(p, d, q)$ , tj.

$$y_t = \frac{\omega_h(B)B^b}{\delta_r(B)} x_t + \frac{\theta_q(B)}{\phi_p(B)} \varepsilon_t \quad (1.3.54)$$

kde řádem  $\omega_h(B)$ ,  $\delta_r(B)$ ,  $\theta_q(B)$  a  $\phi_p(B)$  jsou hodnoty  $h$ ,  $r$ ,  $q$  a  $p$ . Zpoždění pro vstupní řadu  $x_t$  je rovno  $b$ . Hodnoty řad v modelu nyní značíme malými písmeny, neboť předpokládáme, že se jedná o již upravené původní řady  $Y_t$  a  $X_t$ , a že tyto původní řady byly upraveny vzhledem ke stacionaritě. To znamená, že již byly provedeny příslušné transformace k dosažení stacionarity.

Budeme-li odhadovat parametry v tomto modelu, budeme muset vyřešit několik problémů. Vzhledem k tomu, že parametry budeme odhadovat pomocí některého statistického software (SAS, SCA...), musíme zadat vstupní odhady parametrů, které tyto programy požadují. Odhad parametrů polynomů  $\omega_h(B)$  a  $\delta_r(B)$  lze vypočítat na základě předběžných odhadů parametrů  $v_0, v_1, \dots, v_K$  při aplikaci metody LTF. Jako vstupní odhad parametrů pro parametry ARIMA modelu se v praxi často osvědčuje hodnota 0, 1.

Dále budeme řešit problém počátečních vstupních hodnot pro časové proměnné v modelu (1.3.54). Podobný problém se řeší při odhadu parametrů ARIMA modelu při použití podmíněné a nepodmíněné metody nejmenších čtverců. Přepíšme část modelu (1.3.54) s přenosovou funkcí s dosazenými předběžnými odhady do tvaru

$$\hat{y}_t = \frac{\hat{\omega}_h(B)B^{\hat{b}}}{\hat{\delta}_r(B)} x_t = \hat{\delta}_1 \hat{y}_{t-1} + \dots + \hat{\delta}_r \hat{y}_{t-r} + \hat{\omega}_0 x_{t-b} + \dots + \hat{\omega}_h x_{t-h} \quad (1.3.55)$$

Toto je tvar přenosové funkce, který je samozřejmě podmíněn vstupními hodnotami parametrů. Rovnice (1.3.55) nám umožňuje rekurzivní výpočet hodnot  $y_t$  ovšem s tím, že potřebuje vstupní hodnoty pro tento výpočet. Těchto vstupních hodnot je třeba  $k$ , kde  $k = \max(r+1, b+h+1)$ . Volba konstanty  $k$  zohledňuje dostupnost příslušných minulých pozorování  $x_t$  a  $y_t$ .

Abychom mohli pracovat s výrazem (1.3.55), potřebujeme znát počáteční vstupní hodnoty pro  $\hat{y}_t$ ,  $t = 1, 2, \dots$ . Autoři Box a Jenkins (viz [14]) a Wall (viz [85]) navrhli položit tyto startovací hodnoty rovny nule. Liu (viz [58]) však ukázal, že takové hodnoty vedou k vychýlení odhadů parametrů MA modelu. Volba těchto počátečních hodnot bývá tedy řešena různě. Většinou jsme odkázáni na takové počáteční vstupní hodnoty, které jsou aplikovány v námi používaném programu. Systém SCA používá počáteční hodnoty, které

navrhl Liu (viz [58]) (metoda je označována jako „short-cut method treatment“). Tato metoda volby počátečních hodnot dává stabilnější a vydatnější odhady, než metoda původně navržená.

Dalším krokem v procesu odhadu parametrů modelu je výpočet odhadnutých hodnot parametrů poruchové řady

$$\hat{n}_t = y_t - \hat{y}_t, \quad (1.3.56)$$

kde hodnota  $\hat{n}_t = \Delta_L^D \Delta^d \hat{N}_t$  může být vypočítána až pro  $t = k$ . Dále budeme rekurzivně počítat

$$\hat{\varepsilon}_t = \left[ \frac{\hat{\phi}(B)}{\hat{\theta}(B)} \right] \hat{n}_t = \hat{n}_t - \hat{\phi}_1 \hat{n}_{t-1} - \dots - \hat{\phi}_p \hat{n}_{t-p} + \hat{\theta}_1 \hat{\varepsilon}_{t-1} + \dots + \hat{\theta}_q \hat{\varepsilon}_{t-q} \quad (1.3.57)$$

Z (1.3.57) je vidět, že k rekurzivnímu výpočtu  $\hat{\varepsilon}_t$  potřebujeme znát  $q$  předchozích hodnot odhadnutých reziduí (hodnoty  $\hat{\varepsilon}_{t-1}, \dots, \hat{\varepsilon}_{t-q}$ ). Hodnoty  $\hat{\varepsilon}_t$  nemohou být spočítány pro  $t < k + p$ , protože k výpočtu potřebné hodnoty  $\hat{n}_t$  nejsou k dispozici.

Počáteční hodnoty reziduí navrhuje Box a Jenkins (viz [14]) získat pomocí tzv. podmíněné věrohodnostní funkce (conditional likelihood function), zatímco Hillmer a Tiao (viz [48]) navrhují tzv. přesnou věrohodnostní funkci (exact likelihood function). Podmíněná metoda v programu SCA předpokládá, že  $q$  počátečních hodnot reziduí je rovno nule ( $\hat{\varepsilon}_p = \dots = \hat{\varepsilon}_{p-q+1} = 0$ ).

Dalším krokem v procesu odhadu je minimalizace reziduálního součtu čtverců

$$S_{\hat{\varepsilon}}^2 = \sum_{t=k+p}^n \hat{\varepsilon}_t^2 \quad (1.3.58)$$

Tento výpočet lze provést obdobně jako v případě ARIMA modelů. Poté opakujeme všechny kroky procesu odhadu – nahradíme předchozí odhady koeficientů novými odhady, získanými v předchozích krocích, vypočítáme nové hodnoty parametrů přenosové funkce, nové odhady poruchové řady a řady reziduí a nakonec nový součet čtverců reziduí. Tyto kroky opakujeme tak dlouho, dokud není nalezena taková kombinace koeficientů, která minimalizuje výraz (1.3.58).

### 1.3.5. Ověřování modelu

Poté, co je navržen vhodný model a jsou odhadnuty jeho parametry, nastupuje fáze ověřování tohoto modelu. Vzhledem k tomu, že se pracuje s dynamickým regresním modelem, musí se celá fáze ověřování adekvátnosti modelu rozdělit na dvě části podobně jako když se celý model konstruuje. Jednak se ověřuje adekvátnost racionální přenosové funkce, jednak adekvátnost zkonstruovaného ARIMA modelu pro poruchovou řadu  $N_t$ . Ověří se také platnost celého modelu. Pokud by se ukázalo, že model v některém ohledu nevyhovuje, bude třeba takový model přestavět, znovu odhadnout parametry a znovu celý model podrobit kontrole. Kontrola v žádném případě neodhalí, zda byla do modelu zařazena správná vstupní řada (či několik vstupních řad). Volba nejvhodnější vstupní řady, příp. několika vstupních řad, není pouze úkolem statistika, ale většinou to předpokládá spolupráci s odborníky na danou problematiku.

V kapitole 1.3.5.1 uvedeme postup, který lze při ověřování adekvátnosti modelu použít bez ohledu na metodu použitou při identifikaci modelu. Tento postup byl vytvořen především pro identifikaci pomocí LTF, ale je na tolik obecný, že jej lze použít i pro jiné metody identifikace.

#### 1.3.5.1. Obecná metoda ověření modelu

Jeden z hlavních ověřovacích nástrojů adekvátnosti modelu je založen na vzájemné korelační funkci reziduí. Vzájemná korelace se počítá mezi rezidui dynamického regresního modelu v čase  $t$  ( $\hat{\varepsilon}_t$ ) na jedné straně a rezidui ARIMA modelu pro vstupní řadu ( $\hat{\alpha}_t, \hat{\alpha}_{t-1}, \hat{\alpha}_{t-2}, \dots$ ) ve stejném čase  $t$  a v časových posunech  $t-1, t-2, \dots$  na straně druhé. Můžeme přistoupit i k celkovému F-testu, tzn. že budeme testovat všechny reziduální CCF koeficienty najednou. Další částí ověření dynamického regresního modelu je kontrola adekvátnosti ARIMA modelu pro poruchovou řadu  $N_t$ . Postupuje se přitom obvyklým způsobem – testuje se autokorelační funkce reziduí. Je samozřejmě možné pro testování použít i parciální autokorelační funkci. Je možné opět testovat všechny autokorelační funkce reziduí najednou.

Celý proces ověřování je tedy rozdělen do dvou následujících kroků.

1. *Ověření adekvátnosti racionální přenosové funkce.* Adekvátnost racionální přenosové funkce bude ověřována pomocí výběrové vzájemné korelační funkce, která je definována jako

$$r_k^* = \frac{c_{\varepsilon\alpha}(k)}{\hat{\sigma}_\varepsilon \hat{\sigma}_\alpha}, \quad k = 1, 2, 3, \dots \quad (1.3.59)$$

kde  $\hat{\sigma}_\alpha$  a  $\hat{\sigma}_\varepsilon$  jsou výběrové směrodatné odchylky reziduí  $\hat{\alpha}_t$  a  $\hat{\varepsilon}_t$ ,  $c_{\varepsilon\alpha}(k)$  je výběrová kovariance těchto dvou reziduálních řad. Vzhledem k tomu, že se pracuje s výběrovými hodnotami, každý reziduální CCF koeficient je pouze odhadem skutečného odpovídajícího reziduálního CCF koeficientu  $\rho_k^*$ . Ke zjištění významnosti reziduálních CCF koeficientů, je potřeba je podrobit testování. Nulová hypotéza bude mít tvar

$$H_0 : \rho_k^* = 0 \quad (1.3.60)$$

pro  $k = 1, 2, 3, \dots, K$  proti alternativě

$$H_1 : \rho_k^* \neq 0. \quad (1.3.61)$$

Pokud jsou  $\hat{\alpha}_t$  a  $\hat{\varepsilon}_t$  nekorelované a normálně rozdělené a alespoň jedna z řad je bílý šum, potom se dá aproximovat směrodatná odchylka  $r_k^*$  výrazem

$$\sigma(r_k^*) = \frac{1}{\sqrt{n}}, \quad (1.3.62)$$

kde  $n$  je menší z obou rozsahů výběrů řad reziduí  $\hat{\alpha}_t$  a  $\hat{\varepsilon}_t$  (viz [14]). Jakmile některá z hodnot překročí v absolutní hodnotě dvojnásobek směrodatné odchylky (1.3.62), zamítne se nulová hypotéza a hodnota reziduálního CCF koeficientu v tomto bodě je považována za významně se odlišující od nuly.

Místo výše uvedeného postupu se může aplikovat celkový F-test, tzn. testuje se všech  $K$  reziduálních CCF koeficientů najednou. V tomto případě má nulová hypotéza tvar

$$H_0 : \rho_0^* = \rho_1^* = \dots = \rho_K^* = 0 \quad (1.3.63)$$

a stojí proti alternativě

$$H_1 : \text{alespoň jeden korelační koeficient je nenulový.}$$

Testová statistika má tvar (viz [59])

$$S^* = n^2 \sum_{k=0}^K \frac{(r_k^*)^2}{n-k}, \quad (1.3.64)$$

kde  $n$  je menší z obou rozsahů výběrů řad reziduí  $\hat{\alpha}_t$  a  $\hat{\varepsilon}_t$ . Statistika  $S^*$  má za platnosti nulové hypotézy přibližně  $\chi^2$  rozdělení s  $K+I-m$  stupni volnosti, kde  $m$  je počet odhadovaných parametrů přenosové funkce v dynamickém regresním modelu. Hodnota  $m$  nezahrnuje počet parametrů ARIMA modelu pro poruchovou řadu.

Pokud se v předchozích testech prokáže neadekvátnost modelu, je třeba model přestavět. Změna modelu by se měla řídit následujícími pravidly

- primárně je třeba použít reziduální CCF k indikaci neadekvátnosti modelu a teprve v druhé řadě jako nástroj pro výběr alternativního modelu,
- indikuje-li reziduální CCF neadekvátnost přenosové funkce, je třeba se vrátit k hodnotám jednotlivých vah (získaných metodou LTF či libovolnou jinou metodou), znovu posoudit jejich průběh a případně navrhnout alternativní tvar přenosové funkce,
- pokud se po splnění předchozího bodu vhodnost přenosové funkce nezmění k lepšímu, může být užitečné přidat jeden člen do přenosové funkce. Doporučuje se přidat tento člen spíše do čitatele do polynomu  $\omega_h(B)$  (zvýší se tím řád tohoto polynomu) než do jmenovatele, který je tvořen polynomem  $\delta_r(B)$ .

2. *Ověření adekvátnosti ARIMA modelu pro poruchovou řadu  $N_t$ .* Při ověřování adekvátnosti ARIMA modelu pro poruchovou řadu  $N_t$  se testuje autokorelační funkce reziduí  $\hat{\varepsilon}_t$ . Pro testování se může použít i parciální autokorelační funkce, příp. rozšířená autokorelační funkce. Nulová hypotéza má potom tvar

$$H_0 : \rho_k(\varepsilon) = 0 \text{ pro } k = 1, 2, 3, \dots, K, \quad (1.3.65)$$

kde  $\rho_k(\varepsilon)$  je autokorelační funkce reziduí  $\varepsilon_t$ . Nulovou hypotézu testujeme proti alternativě



$$H_1 : \rho_k(\varepsilon) \neq 0. \quad (1.3.66)$$

Testovou statistikou je přímo výběrová autokorelační funkce reziduí  $r_k(\hat{\varepsilon})$ , jejíž hodnoty se opět porovnávají s dvojnásobkem směrodatné odchylky.

Důležitý je fakt, že špatný tvar přenosové funkce může mít za následek i signifikantní hodnoty autokorelačního koeficientu reziduí, ačkoliv uvažovaný ARIMA model je v pořádku. Jestliže obě funkce – jak reziduální CCF tak reziduální ACF indikují neadekvátní model, je dobré nejprve změnit tvar přenosové funkce. Tato změna může totiž způsobit, že model se vylepší natolik, že změny v ARIMA modelu nebudou dále potřebné.

Podobně jako v předchozím případě se mohou testovat všechny autokorelační funkce reziduí  $\hat{\varepsilon}_t$  najednou. Nulová hypotéza

$$H_0 : \rho_1(\varepsilon) = \rho_2(\varepsilon) = \dots = \rho_K(\varepsilon) = 0 \quad (1.3.67)$$

se testuje proti alternativě

$$H_1 : \text{alespoň jeden korelační koeficient je nenulový.}$$

Testová statistika má tvar (viz [59])

$$Q^* = n(n+2) \sum_{k=1}^K \frac{r_k^2(\hat{\varepsilon})}{n-k}. \quad (1.3.68)$$

Za platnosti nulové hypotézy má statistika (3.75) přibližně  $\chi^2$  rozdělení s  $K-m$  stupni volnosti, kde  $m$  je počet odhadovaných parametrů v ARIMA modelu. Hodnota  $m$  tedy nezahrnuje počet parametrů přenosové funkce v dynamickém regresním modelu.

Ve fázi ověřování modelu se jeví jako užitečné znovu zkontrolovat přítomnost odlehlých pozorování v datech. K tomuto účelu se používají např. standardizovaná rezidua. Ty se získají tak, že se hodnoty reziduí  $\hat{\varepsilon}_t$  dělí směrodatnou odchylkou reziduí  $\hat{\sigma}_\varepsilon$ . Za odlehlá pozorování se pak považují ta, kde standardizovaná rezidua jsou větší než číslo 3. Záleží na přísnosti posouzení – při méně striktním posouzení za odlehlá pozorování se považují ta, kde standardizovaná rezidua jsou větší než 3,5 nebo 4. Zjistí-li se přítomnost

odlehlejších pozorování, je třeba pozměnit regresní model začleněním binární vstupní proměnné (viz [16]).

### 1.3.5.2. Obecné zásady výstavby a ověření modelu

Při ověřování modelu se musí mít na zřeteli některé obecné zásady, které je vhodné dodržet. Problematika těchto zásad je diskutována v následujících bodech

- *Jednoduchost modelu.* Dobrý model je takový model, který je co nejjednodušší. To znamená, že už ve fázi výstavby modelu se budou preferovat modely s co nejmenším počtem parametrů. Jednoduchost modelu však nesmí být preferována na úkor jeho kvality.
- *Stabilita, stacionarita a invertibilita.* Dynamický regresní model musí být stabilní. Autoregresní koeficienty v ARIMA modelu pro poruchovou řadu mohou být ověřovány na stacionaritu a koeficienty modelu klouzavých součtů na invertibilitu. Podmínky stacionarity a invertibility jsou popsány v kapitole 1.2.
- *Korelace koeficientů.* V ideálním případě jsou odhadnuté koeficienty v dynamickém regresním modelu nekorelované. Určitá korelace mezi koeficienty vždy existuje. Protože korelaci se zabránit nedá, jde především o to, aby nebyla příliš vysoká, popřípadě aby bylo možno určit, jak velká korelace je vůbec přípustná. Vysoká korelace mezi odhadnutými koeficienty může naznačovat, že model není dostatečně jednoduchý. V takovém případě je dobré důkladně model znovu prostudovat a pokusit se vybrat jiný (i podobný) model s menším počtem parametrů. Toto ovšem nemusí být směřodonné, neboť se může stát, že mezi odhadnutými koeficienty bude existovat poměrně vysoká lineární závislost a další kritéria budou hodnotit model jako přijatelný. To se stává především u koeficientů v ARIMA modelu pro poruchovou řadu.
- *Dynamický regresní model versus ARIMA model.* Porovná-li se pro analyzovaná data reziduální směrodatná odchylka zkonstruovaného dynamického regresního modelu a reziduální směrodatná odchylka vhodného ARIMA modelu zkonstruovaného pro výstupní řadu, může se stát, že ARIMA model vychází z tohoto srovnání lépe. Pak ovšem nemá cenu pracně konstruovat dynamický

regresní model a může se rovnou pro vstupní řadu použít model ARIMA, který vysvětlí chování daných dat lépe. Na druhé straně mohou nastat situace, že se velmi usiluje o konstrukci pomocí dynamického regresního modelu (je potřeba např. zdůraznit závislost výstupní řady na řadě vstupní) a pak se nemusí striktně rozhodovat podle směrodatných odchylek.

- *Přesnost předpovědi.* Mnohdy je nejdůležitějším cílem výstavby vhodného modelu konstrukce předpovědi. To, jak jsou předpovědi přesné, samozřejmě vypovídá o kvalitě modelu. Je tedy důležité kontrolovat přesnost předpovědí, jakmile se objeví nové hodnoty analyzovaných časových řad. Je rovněž možné, že existence nových pozorování si vynutí změnu modelu. V jednodušším případě se nemění tvar modelu jako takový (model je dostatečně stabilní), ale dojde ke změnám v odhadech parametrů. Ve složitějších případech se s výskytem nových dat může objevit požadavek na zařazení další vstupní řady do modelu.

Projde-li model fází ověřování, považuje se za adekvátní a může se použít např. pro tvorbu předpovědí. V případě neadekvátnosti modelu se musí rozlišovat, zda je nesprávně zkonstruovaná přenosová funkce nebo zda je přenosová funkce v pořádku a je pouze nesprávně zkonstruován ARIMA model pro poruchovou řadu. V prvním případě se spustí celý proces identifikace přenosové funkce znovu od začátku a poté se modifikuje model pro poruchovou řadu. Ve druhém případě je situace jednodušší, neboť stačí upravit příslušný ARIMA model pro poruchovou řadu.

### 1.3.6. Předpovědi

Nejdůležitější část při výstavbě modelů pro dané časové řady je tvorba předpovědi. Na základě správných předpovědí se mohou dělat některá důležitá ekonomická či finanční rozhodnutí.

Z technické stránky je výpočet předpovědi v modelech s přenosovou funkcí realizován stejnými procedurami jako odhady parametrů modelu ARIMA. Předpokládejme, že hodnoty parametrů modelu jsou známy, v praxi se tyto hodnoty samozřejmě nahrazují jejich odhady. Dále se bude předpokládat model pouze s jednou

vstupní řadou a pro poruchovou řadu se bude uvažovat ARIMA model bez sezónní složky.

Uvažuje se tedy model

$$Y_t = \frac{\omega(B)}{\delta(B)} X_{t-b} + \frac{\theta(B)}{\phi(B)\Delta^d} \varepsilon_t. \quad (1.3.69)$$

Obě strany rovnice (1.3.69) se vynásobí výrazem  $\delta(B)\phi(B)\Delta^d$ . Poté obdržíme

$$\delta^*(B)Y_t = \omega^*(B)X_{t-b} + \theta^*(B)\varepsilon_t, \quad (1.3.70)$$

kde

$$\begin{aligned} \delta^*(B) &= 1 - \delta_1^*B - \dots - \delta_{p+d+r}^*B^{p+d+r} = \delta(B)\Delta^d\phi(B) \\ \omega^*(B) &= \omega_0^* + \omega_1^*B + \dots + \omega_{p+d+h}^*B^{p+d+h} = \Delta^d\phi(B)\omega(B). \\ \theta^*(B) &= 1 - \theta_1^*B - \dots - \theta_{q+r}^*B^{q+r} = \delta(B)\theta(B) \end{aligned} \quad (1.3.71)$$

Koeficienty polynomů v (1.3.70) označené hvězdičkou se získají porovnáním koeficientů u stejných mocnin  $B$  na obou stranách rovnice (1.3.70).

Předpověď bude konstruována z posledního časového bodu, ve kterém byla napozorována hodnota časové řady, tj. z bodu  $t=n$ . Tento bod se nazývá počátek předpovědi. Předpovědi o  $l$  kroků dopředu ( $l \geq 1$ ) označme  $Y_{n+l}$ . Hodnota  $l$  se nazývá horizont předpovědi. S využitím (1.3.70) se může  $Y_{n+l}$  zapsat jako

$$\begin{aligned} Y_{n+l} &= \delta_1^*Y_{n+l-1} + \delta_2^*Y_{n+l-2} + \dots + \delta_{p+d+r}^*Y_{n+l-p-d-r} + \omega_0^*X_{n+l-b} + \dots + \\ &+ \omega_{p+d+h}^*X_{n+l-b-p-d-h} - \theta_1^*\varepsilon_{n+l-1} - \dots - \theta_{q+r}^*\varepsilon_{n+l-q-r} + \varepsilon_{n+l} \end{aligned} \quad (1.3.71)$$

Předpověď o  $l$  kroků dopředu se konstruuje na základě informací, které jsou k dispozici v počátku předpovědi  $t=n$ . Označme tyto informace obsažené v dostupných datech v čase  $n$  jako  $I_n$  (v čase  $n$  známe pozorování  $(Y_n, Y_{n-1}, \dots; X_n, X_{n-1}, \dots)$ ). Předpověď  $Y_{n+l}$  označme jako  $\hat{Y}_n(l)$ . Tato předpověď je rovna podmíněné střední hodnotě  $Y_{n+l}$  při dané hodnotě  $I_n$ . Pro zjednodušení zápisu budeme tuto podmíněnou střední hodnotu zapisovat do hranatých závorek a nebudeme již uvádět symbol střední hodnoty. Pak lze pro předpověď hodnoty  $Y_{n+l}$  psát

$$\hat{Y}_n(l) = E[Y_{n+l} / I_n] = [Y_{n+l}] = \delta_1^* [Y_{n+l-1}] + \dots + \delta_{p+d+r}^* [Y_{n+l-p-d-r}] + \omega_0^* [X_{n+l-b}] + \dots + \omega_{p+d+h}^* [X_{n+l-b-p-d-h}] - \theta_1^* [\varepsilon_{n+l-1}] - \dots - \theta_{q+r}^* [\varepsilon_{n+l-q-r}] + [\varepsilon_{n+l}] \quad (1.3.72)$$

Hodnoty  $[Y_{n+j}]$  a  $[X_{n+j}]$  jsou pro  $j \leq 0$  napozorované hodnoty časové řady. Obdobně hodnoty  $[\varepsilon_{n+j}]$  jsou pro  $j \leq 0$  odhadnuté hodnoty reziduí  $[\hat{\varepsilon}_{n+j}]$  v modelu s přenosovou funkcí. Naproti tomu hodnoty  $[Y_{n+j}]$  a  $[X_{n+j}]$  jsou pro  $j > 0$  předpovědi  $\hat{Y}_n(j)$  a  $\hat{X}_n(j)$  dané podmíněnými středními hodnotami a  $[\hat{\varepsilon}_{n+j}] = 0$  pro  $j > 0$  - v čase  $n$  nemáme žádné odhady  $\varepsilon_{n+j}$ , takže se nahradí jejich hodnoty očekávanou střední hodnotou a ta je rovna nule.

Odhady  $\hat{X}_n(j)$  se mohou získat různým způsobem. Jednou z možností je zkonstruovat ARIMA model pro  $X_t$ , jinou možností je nahradit tyto hodnoty známými hodnotami – to se může podařit v případě, že tyto hodnoty  $X_t$  jsou dopředu určeny např. jako důsledek politických rozhodnutí, přijetí nových zákonů, změn v daňové soustavě apod. Někdy mohou být hodnoty  $X_t$  dokonce deterministické povahy, lze např. jmenovat proměnné popisující efekt pracovních dnů v časové řadě apod.

Co nás bude asi nejvíce na předpovědi zajímat, bude její přesnost. Za tímto účelem se bude zkoumat rozptyl chyb předpovědi a minimální střední čtvercová chyba předpovědi.

1. *Rozptyl chyb předpovědi.* Abychom mohli vypočítat rozptyl chyb předpovědi, prepíšeme výraz (1.3.69) jako funkci náhodné složky. Nejprve upravíme příslušný ARIMA model pro  $X_t$  do tvaru

$$X_t = \frac{\theta_x(B)}{\Delta^{d_x} \phi_x(B)} \alpha_t. \quad (1.3.73)$$

Výraz (1.3.73) dosadíme za  $X_t$  do (1.3.69), čímž obdržíme

$$Y_t = \eta(B)\alpha_t + \psi(B)\varepsilon_t, \quad (1.3.74)$$

kde

$$\eta(B) = \eta_0 + \eta_1 B + \eta_2 B^2 + \dots = \frac{\omega(B)\theta_x(B)B^b}{\delta(B)\Delta^{d_x}\phi_x(B)} \quad (1.3.75)$$

a

$$\psi(B) = \psi_0 + \psi_1 B + \psi_2 B^2 + \dots = \frac{\theta(B)}{\Delta^d \phi(B)} \quad (1.3.76)$$

s hodnotou  $\psi_0 = 1$ . Oba dva polynomy  $\eta(B)$  a  $\psi(B)$  mohou být teoreticky nekonečně velkého řádu. S využitím definice  $\eta(B)$  můžeme nalézt váhy  $\eta$  porovnáním koeficientů u stejných mocnin  $B$  na obou stranách rovnice

$$\delta(B)\Delta^{d_x}\phi_x(B)\eta(B) = \omega(B)\theta_x(B)B^b. \quad (1.3.77)$$

Analogicky, z definice  $\psi(B)$  získáme váhy  $\psi$  porovnáním koeficientů u stejných mocnin  $B$  na obou stranách rovnice

$$\Delta^d \phi(B)\psi(B) = \theta(B). \quad (1.3.78)$$

S využitím výrazu (1.3.75) zapíšeme budoucí hodnotu  $Y_{n+l}$ , kde  $t = n$  je počátek předpovědi, ve tvaru

$$Y_{n+l} = \eta_0 \alpha_{n+l} + \eta_1 \alpha_{n+l-1} + \eta_2 \alpha_{n+l-2} + \dots + \varepsilon_{n+l} + \psi_1 \varepsilon_{n+l-1} + \psi_2 \varepsilon_{n+l-2} + \dots \quad (1.3.79)$$

Protože předpověď  $Y_{n+l}$  z bodu  $t = n$  není nic jiného než podmíněná střední hodnota, můžeme na základě (1.3.79) psát

$$\hat{Y}_n(l) = [Y_{n+l}] = \eta_l \alpha_n + \eta_{l+1} \alpha_{n-1} + \dots + \psi_l \varepsilon_n + \psi_{l+1} \varepsilon_{n-1} + \dots \quad (1.3.80)$$

Chyba předpovědi s časovým horizontem  $l$  je  $e_n(l) = Y_{n+l} - \hat{Y}_n(l)$ . Odečteme-li (1.3.79) od (1.3.80) obdržíme

$$e_n(l) = \eta_0 \alpha_{n+l} + \eta_1 \alpha_{n+l-1} + \dots + \eta_{l-1} \alpha_{n+1} + \varepsilon_{n+l} + \psi_1 \varepsilon_{n+l-1} + \dots + \psi_{l-1} \varepsilon_{n+1}. \quad (1.3.81)$$

Rozptyl  $e_n(l)$  je potom roven  $\text{var}(e_n(l)) = E[e_n(l)]^2 - (E[e_n(l)])^2$  a jelikož  $E[e_n(l)] = 0$  můžeme psát  $\text{var}(e_n(l)) = E[e_n(l)]^2$ . Výraz (1.3.81) umocníme na druhou a přejdeme ke střední hodnotě, přičemž využijeme skutečnosti, že hodnoty  $\alpha_t$  a  $\varepsilon_t$  jsou vzájemně nezávislé a nejsou autokorelované. Pak obdržíme rozptyl

$$\text{var}(e_n(l)) = \sigma_\alpha^2 \sum_{j=0}^{l-1} \eta_j^2 + \sigma_\varepsilon^2 \sum_{j=0}^{l-1} \psi_j^2. \quad (1.3.82)$$

Pokud jsou navíc  $\alpha_t$  a  $\varepsilon_t$  normálně rozdělené, bude  $e_n(l)$  mít též normální rozdělení, tím pádem určíme interval spolehlivosti pro předpověď  $Y_{n+l}$ . 95%-ní interval spolehlivosti má v takovém případě tvar

$$\left( \hat{Y}_n(l) - 1,96\sqrt{\text{var}(e_n(l))}, \hat{Y}_n(l) + 1,96\sqrt{\text{var}(e_n(l))} \right). \quad (1.3.83)$$

2. **Minimální střední čtvercová chyba.** Spočítáme nyní střední čtvercovou chybu předpovědi  $Y_{n+l}$  (předpověď děláme v čase  $t = n$ ), která je určena rovnicí (1.3.72) a je značena jako  $\hat{Y}_n(l)$ . Za účelem tohoto výpočtu je vhodné vyjádřit budoucí a předpovězené hodnoty ve tvaru náhodné složky. Vyjdeme z výrazu (1.3.80). Můžeme tedy psát

$$\hat{Y}_n(l) = \sum_{j=0}^{\infty} \eta_{l+j}^* \alpha_{n-j} + \sum_{j=0}^{\infty} \psi_{l+j}^* \varepsilon_{n-j}. \quad (1.3.84)$$

Předpověď (1.3.84) odečteme od budoucích hodnot (1.3.79) a po úpravách obdržíme

$$Y_{n+l} - \hat{Y}_n(l) = \sum_{j=0}^{l-1} (\eta_j \alpha_{n+l-j} + \psi_j \varepsilon_{n+l-j}) + \sum_{j=0}^{\infty} ((\eta_{l+j} - \eta_{l+j}^*) \alpha_{n-j} + (\psi_{l+j} - \psi_{l+j}^*) \varepsilon_{n-j}). \quad (1.3.85)$$

Výraz (1.3.85) umocníme na druhou a přejdeme ke střední hodnotě, čímž obdržíme

$$E[Y_{n+l} - \hat{Y}_n(l)]^2 = (\eta_0^2 + \eta_1^2 + \dots + \eta_{l-1}^2) \sigma_\alpha^2 + (1 + \psi_1^2 + \dots + \psi_{l-1}^2) \sigma_\varepsilon^2 + \sum_{j=0}^{\infty} ((\eta_{l+j} - \eta_{l+j}^*)^2 \sigma_\alpha^2 + (\psi_{l+j} - \psi_{l+j}^*)^2 \sigma_\varepsilon^2) \quad (1.3.86)$$

Výraz (1.3.86) nabývá minima pro  $(\eta_{l+j} - \eta_{l+j}^*) = 0$  a  $(\psi_{l+j} - \psi_{l+j}^*) = 0$ .

Z toho plyne, že minimální střední čtvercová chyba předpovědi  $Y_{n+l}$  v čase  $t = n$  je podmíněnou střední hodnotou  $Y_{n+l}$  v časovém bodě  $t = n$ , tzn. že je rovna hodnotě předpovědi v modelu s přenosovou funkcí dané výrazem (1.3.72). Tato hodnota je sama podmíněna hodnotami vstupní a výstupní řady směrem do minulosti.

## 1.4. Modely volatility

V kapitole 1.1. jsme uvedli základní charakteristické vlastnosti výnosů finančních časových řad. Mezi tyto vlastnosti patří i leptokurtické pravděpodobnostní rozdělení a proměnlivá volatilita. Na zvýšení volatility má vliv např. to, že se neobchoduje každý den a akumulace informací se projeví zvýšenou volatilitou v následujících obchodních dnech. Také pravidelné zveřejňování důležitých informací způsobuje její zvýšení. Odlišné reakce volatility na velké kladné a na velké záporné hodnoty míry zisku lze modelovat pomocí modelů volatility. Tyto modely uvažují změny v podmíněných rozptylech (viz [76]). Jejich význam spočívá v tom, že umožňují zachytit měnící se podmínky nejistoty na trhu, což je také v souladu s vývojovými trendy moderní ekonomické teorie. Prostřednictvím nich lze empiricky ověřovat různé ekonomické a finanční teorie týkající se finančního trhu. Lze je využít např. při tvorbě optimálního portfolia či analýze VaR (viz [21]).

Pro modely AR, MA, ARMA (kapitola 1.2) platí, že nepodmíněný rozptyl  $V(y_t)$  je neměnný v čase a podmíněný rozptyl  $V(y_t|y_{t-1}, y_{t-2}, \dots)$  procesu  $\{y_t\}$ . U integrovaných modelů ARIMA je nepodmíněný rozptyl funkcí časové proměnné  $t$  a podmíněný rozptyl je stejně jako v případě stacionárních procesů neměnný v čase. Z uvedeného je patrný problém lineárních modelů při modelování finančních časových řad, podmíněné rozptyly obou typů modelů jsou v čase neměnné. Reálná situace je však často jiná, podmíněné rozptyly jsou v čase proměnlivé. Jednotlivé modely volatility tudíž spočívají ve formulaci vývoje podmíněného rozptylu v čase.

Z hlediska autokorelační struktury časových řad lze modely volatility chápat jako modely nelineární, neboť charakterizují vývoj podmíněného rozptylu stochastického procesu. Zachycují tedy závislosti mezi veličinami stochastického procesu, které nejsou lineární. Z hlediska konkrétní funkční formy modelu podmíněného rozptylu však lze rozlišovat lineární a nelineární modely volatility. Vzhledem k tomu, že dále budeme pracovat pouze s lineárními modely volatility, nebudeme se nelineárními modely volatility v této práci zabývat.



### 1.4.1. Lineární modely volatility - přehled

Modely volatility, kterými se budeme zabývat, vycházejí z představy, že např. autoregresní model AR(1)

$$y_t = \varphi_1 y_{t-1} + \varepsilon_t^* \quad (1.4.1)$$

lze modifikovat do tvaru

$$y_t = \varphi_1 y_{t-1} + \varepsilon_t, \quad (1.4.2)$$

kde  $|\varphi_1| < 1$  a  $\{\varepsilon_t\}$  je podmíněný heteroskedastický proces s podmíněnou střední hodnotou  $E(\varepsilon_t | \Omega_{t-1}) = 0$  a podmíněným rozptylem  $V(\varepsilon_t | \Omega_{t-1}) = E(\varepsilon_t^2 | \Omega_{t-1}) = h_t$ , kde  $\Omega_{t-1}$  je relevantní minulé informace až do času  $t-1$ . Tyto požadavky splňuje model procesu  $\{\varepsilon_t\}$  ve tvaru

$$\varepsilon_t = e_t h_t^{\frac{1}{2}}, \quad (1.4.3)$$

kde veličiny procesu  $\{e_t\}$  jsou nezávislé s nulovou střední hodnotou a jednotkovým rozptylem.

Je-li rozdělení náhodné veličiny  $e_t$  za podmínky informace, která je k dispozici v čase  $t-1$ , normované normální, tj.  $e_t \sim N_{t-1}(0,1)$ , potom je rozdělení náhodné veličiny  $y_t$  za podmínky informace, která je k dispozici v čase  $t-1$ , rovněž normální, avšak s podmíněným rozptylem, měnícím se v závislosti na čase, tj.  $y_t \sim N_{t-1}(0, h_t)$ .

Špičatost nepodmíněného rozdělení  $\varepsilon_t$  je větší nebo rovna špičatosti normovaného normálního rozdělení. Rovnost platí tehdy, jsou-li podmíněné rozptyly konstantní.

Modely volatility poprvé popsal Engle (viz [24]). Jsou charakteristické tím, že podmíněný rozptyl je lineární funkcí veličin  $\varepsilon_{t-1}^2, \varepsilon_{t-2}^2, \dots, \varepsilon_{t-p}^2$ . Vzhledem k tomu, že jde o lineární funkci, označují se tyto modely jako lineární modely volatility.

### 1.4.1.1. Modely ARCH

Nejjednodušší z této skupiny modelů je model ARCH(1).

#### Model ARCH(1)

Tvar podmíněného rozptylu modelu ARCH(1) je následující

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2, \quad (1.4.4)$$

kde

$\varepsilon_t = e_t h_t^{-\frac{1}{2}}$  a veličiny procesu  $\{e_t\}$  jsou nezávislé s nulovou střední hodnotou a jednotkovým rozptylem,

$\alpha_0, \alpha_1$  jsou parametry, pro které platí  $\alpha_0 > 0, \alpha_1 \geq 0$ .

Požadavek  $\alpha_0 > 0, \alpha_1 \geq 0$  na parametry  $\alpha_0$  a  $\alpha_1$  zaručuje, že podmíněný rozptyl je kladné číslo. V případě, že  $\alpha_1 = 0$ , je podmíněný rozptyl v čase konstantní a proces  $\{\varepsilon_t\}$  se označuje za podmíněně homoskedatický.

Autoregresní tvar modelu

$$\varepsilon_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + v_t, \quad (1.4.5)$$

kde  $v_t = h_t(e_t^2 - 1)$ , dostaneme přičtením  $\varepsilon_t^2$  k oběma stranám rovnice (1.4.4) a odečtením  $h_t$ .

Po dosazení  $\varepsilon_t = e_t h_t^{-\frac{1}{2}}$  do (1.4.5) obdržíme

$$\varepsilon_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 h_{t-1} + e_t^2 h_t - h_t, \quad (1.4.6)$$

$$\text{tj. } \varepsilon_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + h_t(e_t^2 - 1). \quad (1.4.7)$$

Podmíněná a nepodmíněná střední hodnota procesu  $\{v_t\}$  je nulová, tzn. že proces není autokorelovaný. Vzhledem k tomu, že jsme získali autoregresní tvar modelu, tak z teorie Box-Jenkinsovy metodologie (kapitola 1.2) vyplývá, že proces ARCH(1) je stacionární v kovariancích pro  $\alpha_1 < 1$ . Nepodmíněný rozptyl procesu  $\{\varepsilon_t\}$  má tvar

$$V(\varepsilon_t) = \frac{\alpha_0}{1 - \alpha_1}, \quad (1.4.8)$$

to znamená, že je konstantní v čase a proces  $\{\varepsilon_t\}$  je nepodmíněně homoskedastický.

Podíváme-li se na vzorec (1.4.7), vidíme, že pokud  $\varepsilon_{t-1}$  je v absolutní hodnotě vysoké, potom bude pravděpodobně i  $\varepsilon_t$  v absolutní hodnotě vysoké. Prostřednictvím modelu ARCH lze tedy zachytit shluky volatility v časové řadě výnosů. Typické pro volatilitu finančních časových řad je právě výskyt segmentů s nízkou a naopak s vysokou volatilitou. Takové segmentování vzniká v důsledku toho, že předchozí vysoká (resp. nízká) volatilita vyvolává s velkou pravděpodobností také vysokou (resp. nízkou) volatilitu v následujícím čase a model ARCH je charakteristický tím, že může zachytit tuto skutečnost.

Model ARCH umožňuje rovněž zachytit vyšší špičatost pravděpodobnostního rozdělení výnosů než má normální rozdělení, tj. 3, což je velmi užitečné, neboť v praxi často dochází k tomu, že rozdělení logaritmů denních, týdenních či čtrnáctidenních výnosů není symetrické a je zešikmené s výraznou špičatostí. To se vyznačuje „tenkým pásem“, kdy část hodnot uvažované náhodné veličiny leží s vysokou pravděpodobností blízko její střední hodnotě.

Zobecněním modelu ARCH(1) je model ARCH(p).

### Model ARCH(p)

Tvar podmíněného rozptylu modelu ARCH(p) je

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_p \varepsilon_{t-p}^2, \quad (1.4.9)$$

kde

$\varepsilon_t = e_t h_t^{\frac{1}{2}}$  a veličiny procesu  $\{e_t\}$  jsou nezávislé s nulovou střední hodnotou a jednotkovým rozptylem,

$\alpha_0, \alpha_1, \dots, \alpha_p$  jsou parametry, pro které platí  $\alpha_0 > 0, \alpha_i \geq 0, i = 1, 2, \dots, p$ .

Podmínka  $\alpha_0 > 0, \alpha_i \geq 0, i = 1, 2, \dots, p$  opět zaručuje, že podmíněný rozptyl bude kladný.

Model ARCH(p) podobně jako model ARCH(1) lze vyjádřit v autoregresním tvaru, tj. ve tvaru modelu AR(p) procesu  $\{\varepsilon_t^2\}$

$$\varepsilon_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_p \varepsilon_{t-p}^2 + v_t, \quad (1.4.10)$$

kde  $v_t = \varepsilon_t^2 - h_t$ . Proces ARCH(p) je stacionární v kovariancích, jestliže kořeny polynomiální rovnice

$$(1 - \alpha_1 B - \alpha_2 B^2 - \dots - \alpha_p B^p) = 0 \quad (1.4.11)$$

leží vně jednotkového kruhu, přičemž  $B$  je operátor zpětného posunutí a rovnici (1.4.10) lze pomocí tohoto operátoru přepsat do tvaru

$$(1 - \alpha_1 B - \alpha_2 B^2 - \dots - \alpha_p B^p) \cdot \varepsilon_t^2 = \alpha_0 + v_t. \quad (1.4.12)$$

Nepodmíněný rozptyl procesu  $\{\varepsilon_t\}$  má tvar

$$V(\varepsilon_t) = \frac{\alpha_0}{1 - (\alpha_1 + \alpha_2 + \dots + \alpha_p)}, \quad (1.4.13)$$

tudíž je konstantní v čase  $t$  a proces  $\{\varepsilon_t\}$  se označuje za nepodmíněně homoskedatický.

#### 1.4.1.2. Modely GARCH

Často se stává, že při modelování finančních časových řad je někdy potřeba použít modely ARCH(p), ve kterých je  $p$  vysoké číslo. V tomto případě musíme odhadovat velké množství parametrů, což je problematické, neboť je na parametry modelu kladena řada omezujících podmínek. Jak je z výše uvedeného zřejmé, jsou to podmínky kladené na nenulovost podmíněného rozptylu a podmínky kladené na parametry z důvodu stacionarity modelu. Tento problém vyřešil Bollerslev (viz [12]), který navrhl rozšířit model ARCH o zpožděný podmíněný rozptyl.

Nejjednodušší případ je samozřejmě rozšíření modelu ARCH(1) o podmíněný rozptyl v prvním zpoždění. V tomto případě mluvíme o modelu GARCH(1,1).

#### Model GARCH(1,1)

Podmíněný rozptyl modelu GARCH(1,1) má tvar

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}, \quad (1.4.14)$$

kde

$\alpha_0 > 0$ ,  $\alpha_1 > 0$  a  $\beta_1 \geq 0$ . Dané podmínky zaručují kladný podmíněný rozptyl.

Model ve tvaru (1.4.14) můžeme pomocí operátoru zpětného posunutí vyjádřit jako

$$(1 - \beta_1 B)h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2. \quad (1.4.15)$$

Jestliže z rovnice (1.4.15) vyjádříme  $h_t$ , obdržíme

$$h_t = (1 - \beta_1 B)^{-1} (\alpha_0 + \alpha_1 \varepsilon_{t-1}^2) = (1 + \beta_1 B + \beta_1^2 B^2 + \dots) \cdot (\alpha_0 + \alpha_1 \varepsilon_{t-1}^2). \quad (1.4.16)$$

Model (1.4.16) není nic jiného než model ARCH( $\infty$ ). Z toho vyplývá, že pokud by bylo vhodné volit model ARCH s mnoha zpožděními, použijeme místo něho model GARCH(1,1).

Stejně jako jsme převedli do autoregresního tvaru model ARCH, převedeme model GARCH(1,1) do tvaru modelu ARMA(1,1), jestliže k oběma stranám rovnice (1.4.14) přičteme  $\varepsilon_t^2$  a odečteme  $h_t$

$$\varepsilon_t^2 = \alpha_0 + (\alpha_1 + \beta_1) \varepsilon_{t-1}^2 + v_t - \beta_1 v_{t-1}, \quad (1.4.17)$$

kde  $v_t = \varepsilon_t^2 - h_t$ . Vzhledem k tomu, že máme nyní model ARMA(1,1), je zřejmé, že model GARCH(1,1) je stacionární v kovariancích, jestliže  $\alpha_1 + \beta_1 < 1$ .

Proces  $\{\varepsilon_t\}$  je opět nepodmíněně homoskedastický, neboť jeho nepodmíněný rozptyl je konstantní v čase. Jeho tvar můžeme vidět níže

$$V(\varepsilon_t) = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}, \quad (1.4.18)$$

Špičatost náhodných veličin  $\varepsilon_t$  je opět větší než špičatost normálního rozdělení, tj.3.

Co se týká hodnot autokorelační a parciální autokorelační funkce procesu  $\{\varepsilon_t^2\}$ , s rostoucím zpožděním exponenciálně klesají. Rychlost poklesu záleží na výši součtu  $\alpha_1 + \beta_1$ . Obecně lze konstatovat, že tvar ACF a PACF procesu  $\{\varepsilon_t^2\}$  odpovídá tvaru těchto funkcí modelu ARMA(p,q).

**Model GARCH(q,p)**

Zobecněním modelů ARCH(p) jsou modely GARCH(q,p). Tyto modely obsahují zpožděný podmíněný rozptyl až do času  $t-q$ . Podmíněný rozptyl obecného modelu GARCH má tvar

$$h_t = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^q \beta_i h_{t-i}. \quad (1.4.19)$$

Pro parametry platí, že  $\alpha_0 > 0, \alpha_i > 0, i = 1, 2, \dots, p, \beta_i \geq 0, i = 1, 2, \dots, q$ . Tyto podmínky zaručují kladný podmíněný rozptyl.

Model GARCH(q,p) je typu ARMA(m,q), kde  $m = \max\{p, q\}$  a dá se přepsat do tvaru

$$\varepsilon_t^2 = \alpha_0 + \sum_{i=1}^m (\alpha_i + \beta_i) \varepsilon_{t-i}^2 - \sum_{i=1}^q \beta_i v_{t-i} + v_t, \quad (1.4.20)$$

kde  $v_t = \varepsilon_t^2 - h_t$ .

Pokud použijeme operátoru zpětného posunutí, můžeme model GARCH(q,p) ve tvaru (1.4.19) přepsat jako

$$h_t = \alpha_0 + \alpha(B) \varepsilon_t^2 + \beta(B) h_t, \quad (1.4.21)$$

kde  $\alpha(B) = \alpha_1 B + \alpha_2 B^2 + \dots + \alpha_p B^p$ ,  $\beta(B) = \beta_1 B + \beta_2 B^2 + \dots + \beta_q B^q$ . (1.4.22)

Pokud od obou stran rovnice (1.4.21) odečteme  $\beta(B) h_t$  a posléze vyjádříme  $h_t$ , obdržíme model GARCH(q,p) ve formě modelu ARCH( $\infty$ )

$$h_t = \frac{\alpha_0}{1 - \beta(B)} + \frac{\alpha(B)}{1 - \beta(B)} \varepsilon_t^2. \quad (1.4.23)$$

Model GARCH(q,p) je stacionární, jestliže leží kořeny polynomiální rovnice

$$1 - \alpha(B) - \beta(B) = 0 \quad (1.4.24)$$

vně jednotkového kruhu.

Vzhledem k tomu, že budeme v této práci využívat pouze modelu ARCH, popř. GARCH, bude následovat už jen stručný přehled lineárních modelů volatility.

### Modely IGARCH

Speciálním případem modelu GARCH(1,1) je model IGARCH(1,1), pro který platí  $\alpha_1 + \beta_1 = 1$  a proces  $\{\varepsilon_t^2\}$  má tvar

$$\varepsilon_t^2 = \alpha_0 + \varepsilon_{t-1}^2 + v_t - \beta_1 v_{t-1}. \quad (1.4.25)$$

Zobecněním modelu IGARCH(1,1) je model IGARCH(q,p), který lze vyjádřit ve tvaru

$$(1 - \alpha(B) - \beta(B))\varepsilon_t^2 = \alpha_0 + (1 - \beta(B))v_t, \quad (1.4.26)$$

kde  $v_t = \varepsilon_t^2 - h_t$  a  $(1 - \alpha(B) - \beta(B)) \equiv \Phi(B)(1 - B)$  a kořeny polynomiální rovnice  $\Phi(B) = 0$  a rovnice  $(1 - \beta(B)) = 0$  leží vně jednotkového kruhu. Model IGARCH(q,p) můžeme zapsat pomocí polynom  $\Phi(B)$  jako

$$\Phi(B)(1 - B)\varepsilon_t^2 = \alpha_0 + (1 - \beta(B))v_t. \quad (1.4.27)$$

### Modely FIGARCH

Pokud se bude autokorelační funkce vyvíjet hyperbolicky a nikoliv exponenciálně jako v předcházejících modelech ARCH, GARCH, IGARCH, můžeme autokorelační strukturu zachytit pomocí frakcionálně integrovaných modelů, které se v tomto případě označují FIGARCH. Obecný model FIGARCH(q,d,p) má formu

$$\Phi(B)(1 - B)^d \varepsilon_t^2 = \alpha_0 + (1 - \beta(B))v_t, \quad (1.4.28)$$

kde  $0 < d < 1$  a kořeny polynomiálních rovnic  $\Phi(B) = 0$  a  $(1 - \beta(B)) = 0$  leží vně jednotkového kruhu.

### 1.4.2. Výstavba modelů volatility

Při výstavbě lineárních a nelineárních modelů volatility lze postupovat následujícím způsobem:

- a) Pro danou časovou řadu se určí vhodný lineární nebo nelineární úrovněvý model.
- b) Otestujeme nulovou hypotézu podmíněné homoskedasticity (tzn. podmíněný rozptyl je v čase neměnný) proti alternativní hypotéze podmíněné heteroskedasticity (tzn. podmíněný rozptyl je proměnlivý) lineárního nebo nelineárního typu. Nás zajímá lineární typ modelu. Test podmíněné heteroskedasticity lineárního typu je založený na principu Lagrangeových multiplikátorů a vychází z formulace modelu ARCH (viz [24]). Vzhledem k tomu, že podmíněný rozptyl modelu ARCH(p) je ve tvaru

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_p \varepsilon_{t-p}^2, \quad (1.4.29)$$

je zřejmé, že je konstantní v případě  $\alpha_1 = \alpha_2 = \dots = \alpha_p = 0$ . Tudíž nulová hypotéza (hypotéza podmíněné homoskedasticity) je ve tvaru

$$H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_p = 0. \quad (1.4.30)$$

Tato nulová hypotéza stojí proti alternativní hypotéze  $H_1$

$H_1$  : alespoň jeden parametr  $\alpha_i, i = 1, 2, \dots, p$  je různý od nuly.

Test bude probíhat následujícím způsobem

- Odhadnou se parametry lineárního nebo nelineárního úrovněvého modelu a získají se rezidua  $\hat{\varepsilon}_t$  a reziduální součet čtverců  $S_{e1}$ .
- Zkonstruuje se regresní model

$$\hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1}^2 + \alpha_2 \hat{\varepsilon}_{t-2}^2 + \dots + \alpha_p \hat{\varepsilon}_{t-p}^2 + u_t \quad (1.4.31)$$

na jehož základě se získá reziduální součet  $S_{e2}$  a index determinace  $R^2$ .



- Testové kritérium je ve tvaru  $T \cdot R^2$ . Za předpokladu platnosti nulové hypotézy má tato testová statistika asymptoticky rozdělení  $\chi_p^2$ , tzn. nulovou hypotézu zamítáme na hladině  $\alpha$ , jestliže  $T \cdot R^2 \geq \chi_p^2(\alpha)$ .
- Pokud máme výběr o malém rozsahu můžeme použít F-verzi výše uvedeného testového kritéria. V tomto případě testová statistika má tvar

$$F = \frac{(S_{e1} - S_{e2})/p}{S_{e2}/(T-p-1)}. \quad (1.4.32)$$

Za platnosti nulové hypotézy lze rozdělení statistiky  $F$  aproximovat rozdělením  $F(p, T-p-1)$  (Fisherovo-Snedecorovo rozdělení o  $p$  a  $T-p-1$  stupních volnosti).

Při testování podmíněné homoskedasticity může dojít k zamítnutí nulové hypotézy podmíněné homoskedasticity v případě, že se nebere ohled na odlehlá pozorování. Řešením tohoto problému je použití testu heteroskedasticity, který je robustní vzhledem k odlehlým pozorováním. Tento test probíhá analogicky jako výše uvedený test, s tím rozdílem, že parametry úrovnového lineárního modelu se odhadnou váženou metodou nejmenších čtverců a dále se pracuje s váženými rezidui.

- c) Odhadneme parametry zvoleného lineárního nebo nelineárního modelu podmíněné heteroskedasticity.

Vzhledem k tomu, že typický model výnosů finančních časových řad se skládá ze dvou částí: lineární či nelineární model úrovně časové řady a lineární či nelineární model volatility časové řady, lze ho zapsat do tvaru

$$X_t = G(\mathbf{X}_t, \eta) + \varepsilon_t, \quad (1.4.33)$$

kde  $\mathbf{X}_t = (1, X_{t-1}, \dots, X_{t-p})$ ,  $G(\mathbf{X}_t, \eta)$  je skeleton lineárního nebo nelineárního autoregresního modelu s parametry  $\eta$ , který je minimálně dvakrát spojitě derivovatelnou funkcí vzhledem k těmto parametrům a  $\{\varepsilon_t\}$  je proces s nulovou podmíněnou střední hodnotou a podmíněným rozptylem  $h_t$  typu lineárního nebo

nelineárního modelu GARCH s parametry  $\varphi$ , a vektor parametrů kompletního modelu je tedy  $\theta = (\eta', \varphi)$ . Tento vektor je třeba odhadnout. Při odhadu parametrů se použije metoda maximální věrohodnosti.

Stejně jako u testování podmíněné heteroskedasticity mohou odlehlá pozorování zkreslit výsledky odhadu parametrů. Tudíž mnohdy se používají metody odhadu parametrů, které by byly vzhledem k těmto odlehlým pozorováním robustní. Jednou z metod je metoda autorů Francese a Ghijselse (viz [30]), která vychází z metody detekce odlehlých pozorování navrženou Chen a Liu (viz [55]).

- d) Ověříme vhodnost daného modelu diagnostickými testy. V tomto případě testujeme nesystematické složky modelu, linearitu a neměnnost parametrů.

*Testování nesystematické složky.* Modely volatility vychází ze základní formulace tvaru

$$\varepsilon_t = e_t h_t^{\frac{1}{2}}, \quad (1.4.34)$$

kde  $e_t$  jsou nezávislé, s nulovou střední hodnotou a jednotkovým rozptylem a v některých případech se předpokládá, že mají normované normální rozdělení. Pokud je model volatility správně určen, potom by standardizovaná rezidua

$$\hat{e}_t = \hat{\varepsilon}_t h_t^{-\frac{1}{2}} \quad (1.4.35)$$

tyto vlastnosti nesystematické složky měla indikovat. Pro diagnostickou kontrolu lze použít některé testy diagnostické kontroly úrovněového modelu jako je např. test autokorelace a normality.

Lundbergh a Teräsvirta (viz [60]) navrhli test zbývající části ARCH. V tomto testu jde o testování, zda standardizovaná rezidua získaná odhadem modelu GARCH vykazují podmíněnou heteroskedasticitu typu ARCH.

*Testování linearity.* Pro testování lineárního modelu GARCH proti nějakému nelineárnímu modelu volatility lze použít testy typu SB („Sign Bias Test“), NSB („Negative Size Bias“) a PSB („Positive Size Bias“), které navrhli Engle a Ng (viz [26]). Problém těchto testů spočívá v tom, že jsou relativně slabé (viz [38]). Navíc nedávají informaci o tom, jak by nelineární alternativa mohla vypadat. Vzniklo proto

několik dalších testů, jejichž alternativní hypotéza má tvar určitého nelineárního modelu. Hagerud (viz [38]) navrhl testovat lineární model GARCH(1,1) proti modelu QGARCH(1,1) a proti modelu LSTGARCH(1,1).

*Testování neměnnosti parametrů.* Lundbergh a Teräsvirta (viz [60]) navrhli test neměnnosti parametrů, ve kterém nulová hypotéza je model GARCH(1,1) a alternativní hypotézou je model

$$h_t = (\alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}) (1 - F(t; \delta)) + (\zeta + \gamma_1 \varepsilon_{t-1}^2 + \tau_1 h_{t-1}) F(t; \delta), \quad (1.4.36)$$

kde

$$F(t; \delta) = \frac{1}{1 + \exp(-\delta \cdot t)} \quad \text{pro } \delta > 0. \quad (1.4.37)$$

Testovaná hypotéza neměnnosti parametrů je

$$H_0 : \alpha_0 = \zeta, \alpha_1 = \gamma_1, \beta_1 = \tau_1. \quad (1.4.38)$$

Problém je, že parametr  $\delta$  z funkce  $F(t; \delta)$  není při testované hypotéze určen. Proto je funkce  $F(t; \delta)$  nahrazena Taylorovou aproximací prvního řádu a je tak vytvořen umělý model tvaru

$$h_t = (\alpha_0^* + \alpha_1^* \varepsilon_{t-1}^2 + \beta_1^* h_{t-1}) + (\zeta^* t + \gamma_1^* \varepsilon_{t-1}^2 t + \tau_1^* h_{t-1} t). \quad (1.4.39)$$

Testová hypotéza je potom ve tvaru

$$H_0 : \zeta^* = \gamma_1^* = \tau_1^* = 0. \quad (1.4.40)$$

- e) Je-li třeba, model modifikujeme.
- f) Použijeme model pro popisné nebo predikční účely.

Jedním z cílů konstrukce lineárních a nelineárních modelů volatility je jejich využití pro tvorbu předpovědí volatility. Předpovědi můžeme konstruovat buď na základě modelů ARIMA za předpokladu podmíněné heteroskedasticity nebo můžeme provádět výpočet předpovědí podmíněného rozptylu na základě lineárních či nelineárních modelů volatility. Konstrukci předpovědí můžeme najít např. Arlt, Arltová (viz [6]).

## 2. Cíl práce a metodika zpracování

### 2.1. Cíl práce

Analýza finančních časových řad v posledních několika letech udělala značný posun vpřed, přičemž vznikla řada modelů popisující právě finanční časové řady, které popisují chování finančních trhů. Jedním z nich je akciový trh. Klíč k úspěšnému obchodování na akciovém trhu je velmi jasný a jednoduchý: ve správný okamžik akcie levně koupit a ve správný okamžik je draze se ziskem prodat. Problémem je, že, kdy nastane ten správný okamžik, nikdo jistě neví. Proto vznikly akciové analýzy, které se na základě různých faktorů snaží nalézt ty, které tržní cenu akcií nějakým způsobem ovlivňují. Základy akciových analýz byly stručně shrnuty v kapitole 1.1.3.4.

V současné době je aplikace akciových analýz předmětem mnohých diskusí, neboť na jejich základě jsou akcioví analytici schopni předpovídat akciové kurzy jen v asi 53%. Tudíž je stále častěji kladena otázka, zda lze vytvořit matematický model, který by výstižně popisoval chování burzy a na jehož základě by se dalo předpovídat její budoucí chování.

Cenový vývoj na akciovém trhu je jedním z nejdůležitějších předstihových ukazatelů v konjunkturálním výzkumu pro indikaci hospodářských cyklů. Vývoj na akciovém trhu předbíhá skutečný vývoj hrubého domácího produktu přibližně o 6 měsíců. Proto analýza a predikce vývoje akcií je důležitá pro pochopení a prognózy vývoje hospodářského produktu. Vzhledem k tomu, že burzovní indexy koncentrují pohyby cen mnoha akcií do jediného čísla, jsou v popředí zájmu právě ony. V České republice mezi nejsledovanější index patří index PX 50.

Při analýze vývoje akcií se běžně používá kombinace akciových analýz uvedených v kapitole 1.1.3.4. Investorská politika se určuje podle globální fundamentální analýzy, výběr titulů se uskutečňuje na základě mikroekonomické fundamentální analýzy a technická analýza slouží ke správnému načasování obchodu (*Timing*). Nemalou roli hraje i psychologická analýza.

Jak již bylo výše řečeno, objevují se pochybnosti o úspěšnosti těchto analýz. Vzhledem k tomu, že vznikla řada nových efektivních a netradičních postupů a metod modelování finančních časových řad, budou tyto metody využity v předkládané práci při určování modelu, který by co nejlépe popisoval vývoj indexu PX 50 a na jehož základě by se mohly dělat prognózy dalšího vývoje tohoto indexu. Kombinace akciových analýz a metod analýzy finančních časových řad by mohla zvednout úspěšnost prognózování akciových kurzů.

Dílčí cíle předkládané práce jsou

1. Určit vhodný model popisující závislost indexu PX 50 na oborových indexech.
2. Určit vhodný model popisující závislost indexu PX 50 a indexu eurozóny, tj. indexu Dow Jones EURO STOXX 50.
3. Najít nejvhodnější model pro vývoj míry zisku indexu PX 50.
4. Najít jednodenní předpovědi volatility pro míru zisku indexu PX 50 a pro míru zisku indexu Standard&Poor's 500 po teroristickém útoku 11.9.2001 na World Trade Center v New Yorku.
5. Najít jednodenní předpovědi volatility pro míru zisku indexu PX 50, indexu Standard&Poor's 500 a indexu Dow Jones Euro Stoxx 50 pro současné období (je duben 2006).
6. Najít předpovědi roční volatility míry zisku indexu PX 50 za období 1995-2006.
7. Najít předpovědi roční volatility míry zisku indexu Standard&Poor's 500 za období 1995-2006 a porovnat je s předpovědi roční volatility míry zisku indexu PX 50.

## **2.2. Metodika zpracování**

Při analýze indexu PX 50 budou využívány metody modelování finančních časových řad jako jsou modely s přenosovou funkcí (kapitola 1.3), modely volatility (kapitola 1.4) a Box – Jenkinsova metodologie (kapitola 1.2), z které předchozí dva modely vycházejí.

Data potřebná pro statistickou analýzu byla dodána Burzou cenných papírů Praha. Statistická analýza byla prováděna pomocí software SCA (modely s přenosovou funkcí),

Eviews 5 (model GARCH), STATISTICA a MATHEMATICA (hledání regresních modelů).

### 2.2.1. Určení vhodného modelu, který popisuje závislost indexu PX 50 na oborových indexech a modelu popisující závislost indexu PX 50 a indexu eurozóny

Při hledání modelu vyjadřujícího závislost indexu PX 50 a oborových indexů použijeme vícerozměrný model s přenosovou funkcí, přičemž řada indexu PX 50 je výstupní řada a jednotlivé řady oborových indexů jsou vstupními řadami. Modely s přenosovou funkcí použijeme i při výstavbě modelu vyjadřující závislost indexu PX 50 (výstupní řada) a indexu eurozóny - Dow Jones EURO STOXX 50 (vstupní řada). Dále bude aplikována i klasická regresní analýza (viz [5]) a některé modely, které se o ní opírají a pro které byly vytvořeny procedury v programu MATHEMATICA (lomená regrese, model dvou regresních přímk) .

Při hledání vhodného modelu budeme pracovat s modely s přenosovou funkcí typu

$$Y_t = v_0 X_t + v_1 X_{t-1} + v_2 X_{t-2} + \dots + v_n X_{t-n} + N_t, \quad (2.2.1)$$

kde

$Y_t$  je výstupní řada,

$X_t$  je vstupní řada,

$v_k, k = 0, \dots, n$  jsou parametry, nazývané impulsní váhy,

$N_t$  je tzv. poruchová řada – tato řada bývá většinou autokorelována. Modelování poruchové řady je možné pomocí Box – Jenkinsovy metodologie.

Dále budeme pracovat s modelem (2.2.1) rozšířeným o další vstupní řady, tzn. s vícerozměrným modelem. Předpokládejme, že máme  $M$  vstupních řad  $X_{1,t}, X_{2,t}, \dots, X_{M,t}$  (v našem případě máme časové řady oborových indexů), potom model bude mít následující tvar

$$Y_t = \sum_{i=1}^M v_i(B) X_{i,t} + N_t, \quad (2.2.2)$$

kde  $v_i(B) = v_{i,0} + v_{i,1}B + v_{i,2}B^2 + \dots, i = 1, 2, \dots, M,$

$B$  je operátor zpětného posunutí,

$N_t$  je poruchová řada.

Při výstavbě modelu s přenosovou funkcí postupujeme následovně (stručné shrnutí postupu uvedeného v kapitole 1.3.)

1. Před vlastní identifikací provedeme předběžné kroky, a to
  - kontrola dat – zkontrolujeme, zda soubor neobsahuje odlehlá pozorování,
  - stacionarita - provedení kontroly, zda řady  $X_t$  a  $Y_t$  (popř. řady  $X_{1,t}, X_{2,t}, \dots, X_{M,t}$  a  $Y_t$ ) jsou stacionární. Pokud řady nejsou stacionární, provede se běžné, případně sezónní diferencování,
  - stabilizace rozptylu na základě Box-Coxovy transformace,
  - sestavení ARIMA modelů pro vstupní ( $X_t$ ) řadu a výstupní ( $Y_t$ ) řadu,
  - testování přítomnosti zpětné vazby mezi vstupními a výstupními řadami (test je obsažen v programu SCA).
2. Identifikace modelu metodou LTF.
3. Odhad parametrů modelu.
4. Ověření modelu. Proces ověřování je rozdělen do dvou kroků, a to
  - ověření adekvátnosti racionální přenosové funkce – pomocí výběrové vzájemné korelační funkce,
  - ověřování adekvátnosti ARIMA modelu pro poruchovou řadu  $N_t$  - na základě autokorelační funkce reziduí.

Výše popsanou výstavbu modelu provedeme v programu SCA.

### 2.2.2. Určení nejvhodnějšího modelu pro míru zisku indexu PX 50

Pro nalezení modelu míry zisku indexu PX 50 použijeme lineární modely volatility (kapitola 1.4).

Při analýze míry zisku indexu PX 50 budeme využívat kombinace autoregresní model řádu  $m$ -tého s modelem volatility GARCH(1,1), tj.

$$r_t = \varphi_1 \cdot r_{t-1} + \dots + \varphi_m \cdot r_{t-m} + \varepsilon_t, \quad \varepsilon_t = \sqrt{h_t} e_t \quad (2.2.3)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta \cdot h_{t-1} \quad e_t \sim N(0,1)$$

a  $r_t$  je míra zisku indexu PX 50.

Při výstavbě modelu postupujeme v následujících krocích

1. Pro danou časovou řadu se určí vhodný lineární nebo nelineární úrovňový model.
2. Otestuje se nulová hypotéza podmíněné homoskedasticity (tzn. podmíněný rozptyl je v čase neměnný) proti alternativní hypotéze podmíněné heteroskedasticity (tzn. podmíněný rozptyl je proměnlivý).
3. Odhadnou se parametry zvoleného modelu podmíněné heteroskedasticity. Při odhadu parametrů se použije metoda maximální věrohodnosti.
4. Ověří se vhodnost daného modelu diagnostickými testy. Testují se nesystematické složky modelu, linearita a neměnnost parametrů. Pro testování nesystematické složky se používá test autokorelace a normality.

Výše popsanou výstavbu modelu provedeme v programu Eviews 5.

### **2.2.3. Určení jednodenní předpovědi volatility míry zisku a určení předpovědi roční volatility míry zisku**

Při určení předpovědi volatility míry zisku aplikujeme exponenciální vyrovnávání. Tato předpověď je využívána v rámci systému RiskMetrics navrženém bankou J.P.Morgan a volně přístupném na Internetu. Hlavním důvodem využití předpovědi volatility pomocí exponenciálního vyrovnávání je, že rutinní použití modelů typu GARCH může být komplikované. Princip exponenciálního vyrovnávání spočívá v tom, že vyhlazená hodnota nebo předpověď se konstruuje jako vážený průměr předchozích hodnot časové řady, přičemž váhy tohoto váženého průměru exponenciálně klesají směrem do minulosti. Při předpovídání volatility má exponenciální vyrovnávání tvar



$$h_t = (1 - \lambda) \sum_{i=0}^{\infty} \lambda^i r_{t-i}^2, \quad (2.2.4)$$

kde  $h_t$  je podmíněný rozptyl,

$r_t$  je míra zisku,

$\lambda$  ( $0 < \lambda < 1$ ) je tzv. vyrovnávací konstanta.

Koeficienty na pravé straně výrazu (2.2.4)  $1 - \lambda, (1 - \lambda)\lambda, (1 - \lambda)\lambda^2, \dots$  představují exponenciálně klesající váhy (jejich součet je roven jedné). Přímé použití tohoto vzorce by však bylo výpočetně náročné, proto se v praxi výhradně používá jeho rekurentní verze tvaru

$$h_t = (1 - \lambda) \cdot r_t^2 + \lambda h_{t-1}. \quad (2.2.5)$$

Na vzorec (2.2.5) pak lze také pohlížet jako na speciální tvar modelu GARCH(1,1), kdy  $\alpha_0 = 0, \alpha_1 = 1 - \lambda, \beta_1 = \lambda$  ( $\alpha_1 + \beta_1 = 1$ , tudíž model je nestacionární).

RiskMetrics používá ve vzorci exponenciálního vyrovnávání (2.2.5) při denních mírách zisku vyrovnávací konstantu  $\lambda = 0,94$  (viz [21]). Speciálně tedy jednodenní předpověď volatility se v RiskMetrics počítá podle vzorce

$$h_t = 0,06 \cdot r_t^2 + 0,94 h_{t-1}. \quad (2.2.6)$$

Pro jednodenní předpověď volatility ve vzorci (2.2.6) potřebujeme počáteční předpověď  $h_0$ . Tato počáteční předpověď bude zkonstruovaná jako

$$h_0 = \sqrt{r_0^2}. \quad (2.2.7)$$

Při předpovědi o více kroků  $T$  vpřed předpovídáme volatilitu časově agregované míry zisku, přičemž v systému RiskMetrics je to dle vztahu

$$h_{t+T} = \sqrt{T} \cdot h_t, \quad (2.2.8)$$

kde  $h_t$  je podle vzorce (2.2.6).

Vzhledem k tomu, že známe vyrovnávací konstantu ( $\lambda = 0,94$ ), lze celou výpočetní proceduru provést v programu Excel.

## 2. Cíl práce a metodika zpracování

### 2.1. Cíl práce

Analýza finančních časových řad v posledních několika letech udělala značný posun vpřed, přičemž vznikla řada modelů popisující právě finanční časové řady, které popisují chování finančních trhů. Jedním z nich je akciový trh. Klíč k úspěšnému obchodování na akciovém trhu je velmi jasný a jednoduchý: ve správný okamžik akcie levně koupit a ve správný okamžik je draze se ziskem prodat. Problémem je, že, kdy nastane ten správný okamžik, nikdo jistě neví. Proto vznikly akciové analýzy, které se na základě různých faktorů snaží nalézt ty, které tržní cenu akcií nějakým způsobem ovlivňují. Základy akciových analýz byly stručně shrnuty v kapitole 1.1.3.4.

V současné době je aplikace akciových analýz předmětem mnohých diskusí, neboť na jejich základě jsou akcioví analytici schopni předpovídat akciové kurzy jen v asi 53%. Tudíž je stále častěji kladena otázka, zda lze vytvořit matematický model, který by výstižně popisoval chování burzy a na jehož základě by se dalo předpovídat její budoucí chování.

Cenový vývoj na akciovém trhu je jedním z nejdůležitějších předstihových ukazatelů v konjunkturálním výzkumu pro indikaci hospodářských cyklů. Vývoj na akciovém trhu předbíhá skutečný vývoj hrubého domácího produktu přibližně o 6 měsíců. Proto analýza a predikce vývoje akcií je důležitá pro pochopení a prognózy vývoje hospodářského produktu. Vzhledem k tomu, že burzovní indexy koncentrují pohyby cen mnoha akcií do jediného čísla, jsou v popředí zájmu právě ony. V České republice mezi nejsledovanější index patří index PX 50.

Při analýze vývoje akcií se běžně používá kombinace akciových analýz uvedených v kapitole 1.1.3.4. Investorská politika se určuje podle globální fundamentální analýzy, výběr titulů se uskutečňuje na základě mikroekonomické fundamentální analýzy a technická analýza slouží ke správnému načasování obchodu (*Timing*). Nemalou roli hraje i psychologická analýza.

Jak již bylo výše řečeno, objevují se pochybnosti o úspěšnosti těchto analýz. Vzhledem k tomu, že vznikla řada nových efektivních a netradičních postupů a metod modelování finančních časových řad, budou tyto metody využity v předkládané práci při určování modelu, který by co nejlépe popisoval vývoj indexu PX 50 a na jehož základě by se mohly dělat prognózy dalšího vývoje tohoto indexu. Kombinace akciových analýz a metod analýzy finančních časových řad by mohla zvednout úspěšnost prognózování akciových kurzů.

Dílčí cíle předkládané práce jsou

1. Určit vhodný model popisující závislost indexu PX 50 na oborových indexech.
2. Určit vhodný model popisující závislost indexu PX 50 a indexu eurozóny, tj. indexu Dow Jones EURO STOXX 50.
3. Najít nejvhodnější model pro vývoj míry zisku indexu PX 50.
4. Najít jednodenní předpovědi volatility pro míru zisku indexu PX 50 a pro míru zisku indexu Standard&Poor's 500 po teroristickém útoku 11.9.2001 na World Trade Center v New Yorku.
5. Najít jednodenní předpovědi volatility pro míru zisku indexu PX 50, indexu Standard&Poor's 500 a indexu Dow Jones Euro Stoxx 50 pro současné období (je duben 2006).
6. Najít předpovědi roční volatility míry zisku indexu PX 50 za období 1995-2006.
7. Najít předpovědi roční volatility míry zisku indexu Standard&Poor's 500 za období 1995-2006 a porovnat je s předpovědi roční volatility míry zisku indexu PX 50.

## **2.2. Metodika zpracování**

Při analýze indexu PX 50 budou využívány metody modelování finančních časových řad jako jsou modely s přenosovou funkcí (kapitola 1.3), modely volatility (kapitola 1.4) a Box – Jenkinsova metodologie (kapitola 1.2), z které předchozí dva modely vycházejí.

Data potřebná pro statistickou analýzu byla dodána Burzou cenných papírů Praha. Statistická analýza byla prováděna pomocí software SCA (modely s přenosovou funkcí),

Eviews 5 (model GARCH), STATISTICA a MATHEMATICA (hledání regresních modelů).

### 2.2.1. Určení vhodného modelu, který popisuje závislost indexu PX 50 na oborových indexech a modelu popisující závislost indexu PX 50 a indexu eurozóny

Při hledání modelu vyjadřujícího závislost indexu PX 50 a oborových indexů použijeme vícerozměrný model s přenosovou funkcí, přičemž řada indexu PX 50 je výstupní řada a jednotlivé řady oborových indexů jsou vstupními řadami. Modely s přenosovou funkcí použijeme i při výstavbě modelu vyjadřující závislost indexu PX 50 (výstupní řada) a indexu eurozóny - Dow Jones EURO STOXX 50 (vstupní řada). Dále bude aplikována i klasická regresní analýza (viz [5]) a některé modely, které se o ní opírají a pro které byly vytvořeny procedury v programu MATHEMATICA (lomená regrese, model dvou regresních přímk) .

Při hledání vhodného modelu budeme pracovat s modely s přenosovou funkcí typu

$$Y_t = v_0 X_t + v_1 X_{t-1} + v_2 X_{t-2} + \dots + v_n X_{t-n} + N_t, \quad (2.2.1)$$

kde

$Y_t$  je výstupní řada,

$X_t$  je vstupní řada,

$v_k, k = 0, \dots, n$  jsou parametry, nazývané impulsní váhy,

$N_t$  je tzv. poruchová řada – tato řada bývá většinou autokorelována. Modelování poruchové řady je možné pomocí Box – Jenkinsovy metodologie.

Dále budeme pracovat s modelem (2.2.1) rozšířeným o další vstupní řady, tzn. s vícerozměrným modelem. Předpokládejme, že máme  $M$  vstupních řad  $X_{1,t}, X_{2,t}, \dots, X_{M,t}$  (v našem případě máme časové řady oborových indexů), potom model bude mít následující tvar

$$Y_t = \sum_{i=1}^M v_i(B) X_{i,t} + N_t, \quad (2.2.2)$$

kde  $v_i(B) = v_{i,0} + v_{i,1}B + v_{i,2}B^2 + \dots, i = 1, 2, \dots, M,$

$B$  je operátor zpětného posunutí,

$N_t$  je poruchová řada.

Při výstavbě modelu s přenosovou funkcí postupujeme následovně (stručné shrnutí postupu uvedeného v kapitole 1.3.)

1. Před vlastní identifikací provedeme předběžné kroky, a to
  - kontrola dat – zkontrolujeme, zda soubor neobsahuje odlehlá pozorování,
  - stacionarita - provedení kontroly, zda řady  $X_t$  a  $Y_t$  (popř. řady  $X_{1,t}, X_{2,t}, \dots, X_{M,t}$  a  $Y_t$ ) jsou stacionární. Pokud řady nejsou stacionární, provede se běžné, případně sezónní diferencování,
  - stabilizace rozptylu na základě Box-Coxovy transformace,
  - sestavení ARIMA modelů pro vstupní ( $X_t$ ) řadu a výstupní ( $Y_t$ ) řadu,
  - testování přítomnosti zpětné vazby mezi vstupními a výstupními řadami (test je obsažen v programu SCA).
2. Identifikace modelu metodou LTF.
3. Odhad parametrů modelu.
4. Ověření modelu. Proces ověřování je rozdělen do dvou kroků, a to
  - ověření adekvátnosti racionální přenosové funkce – pomocí výběrové vzájemné korelační funkce,
  - ověřování adekvátnosti ARIMA modelu pro poruchovou řadu  $N_t$  - na základě autokorelační funkce reziduí.

Výše popsanou výstavbu modelu provedeme v programu SCA.

### 2.2.2. Určení nejvhodnějšího modelu pro míru zisku indexu PX 50

Pro nalezení modelu míry zisku indexu PX 50 použijeme lineární modely volatility (kapitola 1.4).

Při analýze míry zisku indexu PX 50 budeme využívat kombinace autoregresní model řádu  $m$ -tého s modelem volatility GARCH(1,1), tj.

$$r_t = \varphi_1 \cdot r_{t-1} + \dots + \varphi_m \cdot r_{t-m} + \varepsilon_t, \quad \varepsilon_t = \sqrt{h_t} e_t \quad (2.2.3)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta \cdot h_{t-1} \quad e_t \sim N(0,1)$$

a  $r_t$  je míra zisku indexu PX 50.

Při výstavbě modelu postupujeme v následujících krocích

1. Pro danou časovou řadu se určí vhodný lineární nebo nelineární úrovňový model.
2. Otestuje se nulová hypotéza podmíněné homoskedasticity (tzn. podmíněný rozptyl je v čase neměnný) proti alternativní hypotéze podmíněné heteroskedasticity (tzn. podmíněný rozptyl je proměnlivý).
3. Odhadnou se parametry zvoleného modelu podmíněné heteroskedasticity. Při odhadu parametrů se použije metoda maximální věrohodnosti.
4. Ověří se vhodnost daného modelu diagnostickými testy. Testují se nesystematické složky modelu, linearita a neměnnost parametrů. Pro testování nesystematické složky se používá test autokorelace a normality.

Výše popsanou výstavbu modelu provedeme v programu Eviews 5.

### **2.2.3. Určení jednodenní předpovědi volatility míry zisku a určení předpovědi roční volatility míry zisku**

Při určení předpovědi volatility míry zisku aplikujeme exponenciální vyrovnávání. Tato předpověď je využívána v rámci systému RiskMetrics navrženém bankou J.P.Morgan a volně přístupném na Internetu. Hlavním důvodem využití předpovědi volatility pomocí exponenciálního vyrovnávání je, že rutinní použití modelů typu GARCH může být komplikované. Princip exponenciálního vyrovnávání spočívá v tom, že vyhlazená hodnota nebo předpověď se konstruuje jako vážený průměr předchozích hodnot časové řady, přičemž váhy tohoto váženého průměru exponenciálně klesají směrem do minulosti. Při předpovídání volatility má exponenciální vyrovnávání tvar

$$h_t = (1 - \lambda) \sum_{i=0}^{\infty} \lambda^i r_{t-i}^2, \quad (2.2.4)$$

kde  $h_t$  je podmíněný rozptyl,

$r_t$  je míra zisku,

$\lambda$  ( $0 < \lambda < 1$ ) je tzv. vyrovnávací konstanta.

Koeficienty na pravé straně výrazu (2.2.4)  $1 - \lambda, (1 - \lambda)\lambda, (1 - \lambda)\lambda^2, \dots$  představují exponenciálně klesající váhy (jejich součet je roven jedné). Přímé použití tohoto vzorce by však bylo výpočetně náročné, proto se v praxi výhradně používá jeho rekurentní verze tvaru

$$h_t = (1 - \lambda) \cdot r_t^2 + \lambda h_{t-1}. \quad (2.2.5)$$

Na vzorec (2.2.5) pak lze také pohlížet jako na speciální tvar modelu GARCH(1,1), kdy  $\alpha_0 = 0, \alpha_1 = 1 - \lambda, \beta_1 = \lambda$  ( $\alpha_1 + \beta_1 = 1$ , tudíž model je nestacionární).

RiskMetrics používá ve vzorci exponenciálního vyrovnávání (2.2.5) při denních mírách zisku vyrovnávací konstantu  $\lambda = 0,94$  (viz [21]). Speciálně tedy jednodenní předpověď volatility se v RiskMetrics počítá podle vzorce

$$h_t = 0,06 \cdot r_t^2 + 0,94 h_{t-1}. \quad (2.2.6)$$

Pro jednodenní předpověď volatility ve vzorci (2.2.6) potřebujeme počáteční předpověď  $h_0$ . Tato počáteční předpověď bude zkonstruovaná jako

$$h_0 = \sqrt{r_0^2}. \quad (2.2.7)$$

Při předpovědi o více kroků  $T$  vpřed předpovídáme volatilitu časově agregované míry zisku, přičemž v systému RiskMetrics je to dle vztahu

$$h_{t+T} = \sqrt{T} \cdot h_t, \quad (2.2.8)$$

kde  $h_t$  je podle vzorce (2.2.6).

Vzhledem k tomu, že známe vyrovnávací konstantu ( $\lambda = 0,94$ ), lze celou výpočetní proceduru provést v programu Excel.

### 3. Výsledky disertační práce

V předchozích kapitolách jsme se zabývali metodami modelování finančních časových řad, počínaje standardními metodami (kapitola 1.2 – Box-Jenkinsova metodologie) a konče nejmodernějšími přístupy modelování časových řad (kapitola 1.3 – Modely s přenosovou funkcí, kapitola 1.4 – Modely volatility).

V této části se zaměříme na využití těchto metod v praxi, konkrétně se budeme zabývat modelováním indexu PX 50, který do března 2006 reprezentoval akciový trh v České republice. Index PX 50 stál v popředí zájmu nejen široké veřejnosti, ale i odborníků zabývajících se analýzami akciového trhu. Ke zpracování vytyčených cílů (viz kapitola 2.1) týkajících se indexu PX 50 použijeme metod analýzy finančních časových řad.

Jak již bylo řečeno v kapitole 2 (Cíl a metodika), budeme při zpracování užívat software Eviews 5, SCA, STATISTICA a MATHEMATICA. Vzhledem k tomu, že výstupy z těchto programů jsou mnohdy příliš rozsáhlé, nebudeme je (kromě jednoho, viz Příloha - Výstup z programu SCA) celé uvádět. V textu budou obsaženy pouze části výstupů, a to ty, které budou obsahovat důležité kroky konstrukce jednotlivých modelů.

#### 3.1. Index PX 50 versus oborové indexy

Index PX 50 a oborové indexy byly součástí Burzy cenných papírů Praha (BCPP) až do začátku roku 2006. Vyvstává otázka, zda mezi těmito burzovními indexy existovala závislost, konkrétně zda hodnota indexu PX 50 závisela na oborových indexech a zda závislost byla ve stejných časových bodech nebo zda existovala určitá časová prodleva mezi těmito indexy, tj. indexem PX 50 a oborovými indexy. Vzhledem k tomu, jak byly tyto indexy konstruovány (index PX 50 byl podrobně popsán v kapitole 1.1. a charakteristika oborových indexů bude uvedena níže), lze předpokládat kladnou odpověď na otázku závislosti. Závislost se budeme snažit popsat modelem s přenosovou funkcí, který byl podrobně uveden v kapitole 1.3, přičemž předpokládáme, že index PX 50 je



výstupní řadou a oborové indexy jsou vstupními řadami, tzn. budeme uvažovat vícenásobný vstup.

Dříve než přejdeme k vlastní analýze, popíšeme velice stručně oborové indexy. Oborové indexy byly v souladu s burzovní oborovou klasifikací oficiálně publikovány od 6. 4. 1995. Báze jednotlivých indexů byly generovány automaticky ve shodě s indexem PX-GLOB. V případě, že počet bazických emisí byl menší než 3, výpočet indexu byl ukončen. Oddělení statistiky BCPP denně stanovovalo hodnoty oborových indexů a globálního indexu od 30.9.1994, kdy byla výchozí hodnota indexu nastavena na 1 000 bodů. Výpočetní vzorce byly stejné jako u indexu PX 50 (kapitola 1.1.).

V dubnu 1995 zavedla burza globální index PX –GLOB a 19 oborových indexů ve shodě se zavedenou burzovní oborovou klasifikací. Globální index PX –GLOB obsahuje ve své bázi všechny emise akcií včetně investičních fondů a podílových listů registrované na burze, u nichž byl nejpozději v předchozí seanci stanoven závěrečný kurz. V roce 2005 vzrostla hodnota indexu PX –GLOB o 46,9%. Na počátku roku 2006 byla báze indexu PX –GLOB tvořena 39 akciovými emisemi. Analogicky byly automaticky generovány báze oborových indexů. Báze oborového indexu musela obsahovat minimálně 3 emise. Z důvodu nedostatečného počtu bazických emisí byl nejdříve ukončen výpočet 16 oborových indexů k únoru 2005 a zbývající tři indexy byly naposledy propočteny k 28.únoru 2006.

V kapitole 1.1. jsme uvedli, že došlo k určitým změnám na pražské burze – indexy PX 50 a PX-D byly nahrazeny jediným indexem PX. Změny se nevyhnuly ani oborovým indexům pražské burzy. Jak již bylo výše poznamenáno, datum 28.února 2006 je posledním dnem, kdy byly propočteny oborové indexy BI07 – chemický, farmaceutický a gumárenský průmysl, BI12 – energetika a BI16 – služby.

V tabulce 3.1 můžeme najít přehled platných oborových indexů ke dni 28.2.2006 a v tabulce 3.2 je uveden přehled oborových indexů zrušených k únoru 2005 z důvodu nedostatečného počtu bazických emisí, včetně datumu zrušení.

Tabulka 3.1 : Přehled platných oborových indexů k 28.2.2006

Označení indexu	Název oboru
BI07	Chemický, farmaceutický a gumárenský průmysl
BI12	Energetika
BI16	Služby

Tabulka 3.2- Přehled oborových indexů zrušených z důvodu nedostatečného počtu bazických emisí k únoru 2005

Označení indexu	Název oboru	Zrušen ke dni
BI01	Zemědělství	15.2.1999
BI02	Výroba potravin	1.3.2000
BI06	Dřevařský a papírenský průmysl	2.5.2000
BI17	Bižuterie, sklo a keramika	25.9.2001
BI09	Hutnictví, zpracování kovů	21.12.2001
BI19	Ostatní	11.6.2002
BI05	Textilní, oděvní a kožedělný průmysl	29.8.2002
BI18	Investiční fondy	19.9.2002
BI03	Výroba nápojů a tabáku	15.7.2003
BI11	Elektrotechnika a elektronika	23.9.2004
BI10	Strojírenství	30.12.2004
BI13	Doprava a spoje	14.2.2005
BI08	Stavebnictví a průmysl stavebních hmot	29.4.2005
BI14	Obchod	12.7.2005
BI04	Těžba a zpracování nerostů	21.7.2005
BI15	Peněžnictví	1.9.2005

Z tabulky 3.2 vidíme, že oborové indexy od roku 1999 postupně zanikaly, tudíž při analýze vezmeme napozorovaná data od 8.1.1996 do 30.12.1999, tedy data za celé 4 roky existence všech oborových indexů (kromě indexu oboru BI01 – zemědělství). Jedná se o souvislé denní časové řady (denní znamená v tomto případě všechny dny, kdy BCPP obchodovala). Celkem máme k dispozici 1004 pozorování pro každý index. Výjimku tvoří oborový index BI01 (zemědělství), který přestal být od 15.2.1999 zveřejňován, tudíž časová řada tohoto indexu je kratší než uvažované čtyři roky a do dalšího zpracování ji nezahrneme. Data bereme od 8.1.1996, i když oborové indexy jsou stanovovány od 30.9.1994. Předpokládejme, že v období od 30.9.1994 do 8.1.1996 došlo k určité „normalizaci“ obchodování na burze, které do té doby v České republice neexistovalo.

Vzhledem k tomu, jak byla konstruována báze indexu PX 50, je odůvodněné se domnívat, že mezi tímto indexem a indexy oborovými bude existovat určitá závislost, kterou bude možné popsat modelem s přenosovou funkcí. Zvolíme tedy model, kde jako vstupy budeme uvažovat 18 oborových indexů BI02,..., BI19 a výstupem bude index PX 50. Tento model je poměrně složitý a hlavně náročný na zpracování. Z tohoto důvodu budeme zkoumat závislost mezi výstupem a vstupními řadami pouze ve stejném časovém bodě a v čase o jedno období dozadu. Dá se však předpokládat, že některé oborové indexy ze zpracování vypadnou, neboť se ukáže, že do modelu žádnou podstatnou informaci nepřináší. Navíc se po prvním výpočtu (na základě studia reziduí modelu) ukazuje, že všechny řady je třeba diferencovat a konstruovat model pro první diference. Pro identifikaci modelu program SCA používá metodu LTF, která byla podrobně popsána v kapitole 1.3. Na základě této metody se zjistilo (viz níže uvedený výstup z programu SCA), že nejde o žádné časové zpoždění mezi indexem PX 50 a oborovými indexy, tzn. závislost je ve stejném časovém okamžiku. Tím se nám úloha redukuje na regresi mezi indexem PX 50 a oborovými indexy. Významné hodnoty koeficientů jsou zvýrazněny ve výstupu SCA tmavším stínováním.

## Výstup z programu SCA

PARAMETR LABEL	VARIABLE NAME	NUM./ DENOM.	FACTOR	ORDER	CONS- TRAI NT	VALUE	STD ERROR	T VALUE	
1	CONST	CNST	1	0	NONE	.0135	.0288	.47	
2	B2_0	B2BC	NUM.	1	0	NONE	.0400	.0033	12.30
3	B2_1	B2BC	NUM.	1	1	NONE	.0043	.0032	1.33
4	B3_0	B3BC	NUM.	1	0	NONE	.0176	.0021	8.46
5	B3_1	B3BC	NUM.	1	1	NONE	.0017	.0021	.83
6	B4_0	B4BC	NUM.	1	0	NONE	.0171	.0028	6.21
7	B4_1	B4BC	NUM.	1	1	NONE	-.0010	.0028	-.37
8	B5_0	B5BC	NUM.	1	0	NONE	.0026	.0072	.37
9	B5_1	B5BC	NUM.	1	1	NONE	-.0146	.0071	-2.04
10	B6_0	B6BC	NUM.	1	0	NONE	.0174	.0019	8.91
11	B6_1	B6BC	NUM.	1	1	NONE	.0015	.0019	.78
12	B7_0	B7BC	NUM.	1	0	NONE	.0978	.0080	12.27
13	B7_1	B7BC	NUM.	1	1	NONE	-.0119	.0079	-1.51
14	B8_0	B8BC	NUM.	1	0	NONE	.0267	.0053	5.05
15	B8_1	B8BC	NUM.	1	1	NONE	-.0067	.0052	-1.29
16	B9_0	B9BC	NUM.	1	0	NONE	.0176	.0040	4.36
17	B9_1	B9BC	NUM.	1	1	NONE	-.0008	.0040	-.19
18	B10_0	B10BC	NUM.	1	0	NONE	.0462	.0045	10.22
19	B10_1	B10BC	NUM.	1	1	NONE	-.0080	.0044	-1.82
20	B11_0	B11BC	NUM.	1	0	NONE	-.0023	.0067	-.35
21	B11_1	B11BC	NUM.	1	1	NONE	-.0015	.0067	-.23
22	B12_0	B12BC	NUM.	1	0	NONE	.2140	.0035	61.88
23	B12_1	B12BC	NUM.	1	1	NONE	-.0052	.0035	-1.48
24	B13_0	B13BC	NUM.	1	0	NONE	.1170	.0017	68.26
25	B13_1	B13BC	NUM.	1	1	NONE	-.0008	.0017	-.46
26	B14_0	B14BC	NUM.	1	0	NONE	-.0130	.0062	-2.09
27	B14_1	B14BC	NUM.	1	1	NONE	.0023	.0063	.36
28	B15_0	B15BC	NUM.	1	0	NONE	.1562	.0031	50.99
29	B15_1	B15BC	NUM.	1	1	NONE	.0050	.0031	1.63
30	B16_0	B16BC	NUM.	1	0	NONE	.0218	.0045	4.82
31	B16_1	B16BC	NUM.	1	1	NONE	-.0088	.0045	-1.96
32	B17_0	B17BC	NUM.	1	0	NONE	.0321	.0024	11.68
33	B17_1	B17BC	NUM.	1	1	NONE	-.0026	.0027	-.95
34	B18_0	B18BC	NUM.	1	0	NONE	.0005	.0052	.10
35	B18_1	B18BC	NUM.	1	1	NONE	.0011	.0052	.21
36	B19_0	B19BC	NUM.	1	0	NONE	-.0067	.0038	-1.78
37	B19_1	B19BC	NUM.	1	1	NONE	.0004	.0037	-.11
38	PX	D-AR	1	1	NONE	.1323	.0441	3.00	
EFFECTIVE NUMBER OF OBSERVATIONS . . . . .					532				
R-SQUARE . . . . .					1.000				
RESIDUAL STANDARD ERROR . . . . .					.511560E+00				

Významné jsou hodnoty koeficientů u oborových indexů BI02 (výroba potravin), BI03 (výroba nápojů a tabáku), BI06 (dřevařský a papírenský průmysl), BI07 (chemický, farmaceutický a gumárenský průmysl), BI08 (stavebnictví a průmysl stavebních hmot),

BI09 (hutnictví, zpracování kovů), BI10 (strojírenství), BI12 (energetika), BI13 (doprava a spoje), BI15 (peněžnictví), BI16 (služby) a BI17 (bižuterie, sklo a keramika).

Sestrojíme tedy regresní model, kde vysvětlovanou proměnnou je řada prvních diferencí indexu PX 50 a vysvětlujícími proměnnými jsou první difference oborových indexů (byl zkoušen i model bez diferencí, ale průběh reziduálních ACF jednoznačně naznačoval nutnost diferencování). Na základě dílčích t-testů a podmíněného rozkladu součtu čtverců modelu na přínos jednotlivých proměnných je zřejmé, že některé regresní koeficienty se jeví jako nevýznamné a proto je nutné příslušné proměnné z modelu vypustit, stejně jako příslušnou konstantu. Po úpravě modelu obdržíme konečné odhady (níže je uvedena pouze část regresního výstupu s odhadem koeficientů):

#### Výstup z programu SCA

REGRESSION ANALYSIS FOR THE VARIABLE				PX
PREDICTOR	COEFFICIENT	STD.ERROR	T-VALUE	
B2BC	.03299	.00332	9.94	
B3BC	.00941	.00211	4.47	
B4BC	.02479	.00254	9.75	
B6BC	.01526	.00248	6.16	
B7BC	.13568	.00893	15.19	
B9BC	.02512	.00499	5.04	
B12BC	.22487	.00420	53.56	
B13BC	.12785	.00123	103.90	
B15BC	.15111	.00346	43.72	
B16BC	.02317	.00500	4.64	
B17BC	.03070	.00362	8.48	

CORRELATION MATRIX OF REGRESSION COEFFITIENS								
B2BC	1.00							
B3BC	-.09	1.00						
B4BC	-.08	-.06	1.00					
B6BC	.03	-.02	-.03	1.00				
B7BC	-.06	-.04	-.08	-.07	1.00			
B9BC	-.05	-.06	-.23	-.08	-.14	1.00		
B12BC	.492E-02	-.10	-.03	.02	-.15	-.04	1.00	
B13BC	.643E-02	-.38E-03	.06	.618E-02	.19-	.49E-02	-.04	1.00
B15BC	-.04	.478E-02	-.06	.206E-02	-.24	-.03	-.03	-.04
B16BC	-.02	-.03	-.03	-.05	.02	-.09	-.03	-.03
B17BC	-.97E-02	-.02	-.07	.582E-02	-.98E-02	-.05	-.05	-.05
B2BC		B3BC	B4BC	B6BC	B7BC	B9BC	B12	

B13BC	1.00			
B15BC	-.27	1.00		
B16BC	-.03	-.06	1.00	
B17BC	.02	.03	-.04	1.00
	B13BC	B15BC	B16BC	B17BC
S =	.7993	R**2= 97.8%	R**2 (ADJ) = 97.8%	
-----				
ANALYSIS OF VARIANCE TABLE				
-----				
SOURCE	SUM OF SQUARES	DF	MEAN SQUARE	F-RATIO
REGRESSION	28266.686	11	2569.699	4022.287
RESIDUAL	633.754	992	.639	
ADJ. TOTAL	28900.440	1003		
SOURCE	SEQUENTIAL SS	DF	MEAN SQUARE	F-RATIO
B2BC	689.238	1	689.238	1078.848
B3BC	703.556	1	703.556	1101.258
B4BC	738.036	1	738.036	1155.231
B6BC	126.752	1	126.752	198.402
B7BC	7606.076	1	7606.076	11905.605
B9BC	168.015	1	168.015	262.990
B12BC	7701.154	1	7701.154	12054.428
B13BC	9246.109	1	9246.109	4472.710
B15BC	1225.817	1	1225.817	1918.742
B16BC	15.951	1	15.951	24.968
B17BC	45.979	1	45.979	71.971

Nakonec obdržíme model (všechny proměnné v modelu jsou již diferencované):

$$\begin{aligned}
 PX_t = & 0,033 \cdot BI02_t + 0,009 \cdot BI03_t + 0,025 \cdot BI04_t + 0,015 \cdot BI06_t + 0,136 \cdot BI07_t + 0,025 \cdot BI09_t + \\
 & + 0,225 \cdot BI12_t + 0,128 \cdot BI13_t + 0,151 \cdot BI15_t + 0,023 \cdot BI16_t + 0,03 \cdot BI17_t,
 \end{aligned}
 \tag{3.1.1}$$

kde směrodatná odchylka reziduí má hodnotu  $\sigma_\varepsilon = 0,7948$ .

Z tabulky korelační matice odhadnutých koeficientů je vidět, že mezi odhady neexistuje ani v jednom případě významná lineární závislost. Rovněž tabulka rozkladů součtů čtverců modelu na přínos jednotlivých proměnných opravňuje zařazení daných proměnných do modelu. Také studium průběhu ACF a PACF potvrdilo, že rezidua tvoří bílý šum. Lze tedy odhadnutý model prohlásit za adekvátní. O kvalitě modelu nakonec svědčí i index determinace, který nabývá hodnoty 0,978.

Podívejme se ještě (bez uvedení příslušných výstupů z programu SCA), jak vypadá ARIMA model indexu PX 50. Nesmíme ovšem zapomenout, že nezpracováváme poslední

data indexu PX 50, ale data od 8.1.1996 do 30.12.1999, tudíž budeme mít model fungující pro toto období. Model ARIMA pro index PX 50 je tedy

$$(1 - B)\left(1 - 0,2007 \cdot B - 0,1187 \cdot B^{10}\right)PX_t = \varepsilon_t . \quad (3.1.2)$$

Model (3.1.2) má reziduální směrodatnou odchylku  $\sigma_\varepsilon = 5,22306$ . Je zřejmé, že se jedná (vzhledem k povaze řady – jde o burzovní index) o poměrně vysoký stupeň autoregrese. Z porovnání směrodatných odchylek regresního modelu a modelu ARIMA vychází podstatně lépe regresní model. Na druhé straně ARIMA model je nesporně jednodušším modelem s daleko menším počtem odhadovaných parametrů. Vzhledem k tomu, že cílem naší analýzy bylo vytvořit model zkoumající vztah mezi indexem PX 50 a oborovými indexy, dáme přednost modelu (3.1.1).

Vraťme se k modelu (3.1.1). V tabulce 3.3 jsou shrnuty oborové indexy s daty jejich zrušení a hodnotou příslušných regresních koeficientů odpovídajících tomuto modelu. Po shlédnutí hodnot příslušných regresních koeficientů, můžeme konstatovat: Pokud by střední hodnoty jednotlivých oborových indexů byly stejné, tak bychom mohli říci, že na vysvětlení indexu PX 50 se nejvíce podílí oborové indexy, které mají nejvyšší regresní koeficienty, tj. BI12 –energetika (regresní koeficient 0,225), BI15 – peněžnictví (regresní koeficient 0,151) a BI07 –chemický, farmaceutický a gumárenský průmysl (regresní koeficient 0,136), tedy oborové indexy, které patřily mezi poslední zrušené indexy. Na základě statistického testu se však zjistilo, že střední hodnoty oborových indexů stejné nejsou, tudíž tento závěr udělat nemůžeme.

Vzhledem k tomu, že nás zajímá, který z oborových indexů se největší měrou podílí na vysvětlení indexu PX 50, provedeme na závěr metodu, která postupně zařazuje do modelu proměnné, které postupně nejvíce vysvětlují vysvětlovanou proměnnou, tj. v našem případě index PX 50. Aplikaci této metody uděláme v programu STATISTICA. Výstup můžeme vidět v tabulce 3.4.

Tabulka 3.3 – Oborové indexy a hodnoty jejich regresních koeficientů

Označení indexu	Název oboru	Zrušen ke dni	Zastoupení v regresním modelu (udáno hodnotou statisticky významného regresního koeficientu)
BI01	Zemědělství	15.2.1999	neuvažováno
BI02	Výroba potravin	1.3.2000	0,033
BI06	Dřevařský a papírenský průmysl	2.5.2000	0,015
BI17	Bižuterie, sklo a keramika	25.9.2001	0,03
BI09	Hutnictví, zpracování kovů	21.12.2001	0,025
BI19	Ostatní	11.6.2002	Statisticky nevýznamný
BI05	Textilní, oděvní a kožedělný průmysl	29.8.2002	Statisticky nevýznamný
BI18	Investiční fondy	19.9.2002	Statisticky nevýznamný
BI03	Výroba nápojů a tabáku	15.7.2003	0,009
BI11	Elektrotechnika a elektronika	23.9.2004	Statisticky nevýznamný
BI10	Strojírenství	30.12.2004	Statisticky nevýznamný
BI13	Doprava a spoje	14.2.2005	0,128
BI08	Stavebnictví a průmysl stavebních hmot	29.4.2005	Statisticky nevýznamný
BI14	Obchod	12.7.2005	Statisticky nevýznamný
BI04	Těžba a zpracování nerostů	21.7.2005	0,025
BI15	Peněžnictví	1.9.2005	0,151
BI07	Chemický, farmaceutický a gumárenský průmysl	28.2.2006	0,136
BI12	Energetika	28.2.2006	0,225
BI16	Služby	28.2.2006	0,023



Tabulka 3.4 – Výstup z programu STATISTICA

	Step	Multiple	Multiple	R-square	F - to	p-level	Variables
<b>BI13</b>	1	0,869455	0,755953	0,755953	3100,661	0,000000	1
<b>BI12</b>	2	0,946767	0,896368	0,140415	1354,941	0,000000	2
<b>BI15</b>	3	0,978810	0,958069	0,061701	1470,023	0,000000	3
<b>BI07</b>	4	0,983290	0,966858	0,008789	264,673	0,000000	4
<b>BI04</b>	5	0,985599	0,971406	0,004548	158,566	0,000000	5
<b>BI02</b>	6	0,986890	0,973952	0,002546	97,368	0,000000	6
<b>BI17</b>	7	0,987841	0,975830	0,001878	77,291	0,000000	7
<b>BI10</b>	8	0,988655	0,977439	0,001609	70,875	0,000000	8
<b>BI06</b>	9	0,989141	0,978400	0,000961	44,201	0,000000	9
<b>BI16</b>	10	0,989400	0,978912	0,000512	24,088	0,000001	10
<b>BI09</b>	11	0,989582	0,979273	0,000361	17,236	0,000036	11
<b>BI03</b>	12	0,989744	0,979594	0,000321	15,586	0,000084	12
<b>BI08</b>	13	0,989817	0,979737	0,000143	6,996	0,008297	13
<b>BI19</b>	14	0,989874	0,979850	0,000112	5,510	0,019108	14
<b>BI18</b>	15	0,989930	0,979962	0,000112	5,536	0,018828	15

Po zhlédnutí tabulky 3.4 vidíme, že na vysvětlení indexu PX 50 se nejvíce podílí oborový index BI13 (doprava a spoje). Dále jsou to indexy BI12 (energetika), BI15 (peněžnictví) a BI07 (chemický, farmaceutický a gumárenský průmysl). Při zahrnutí těchto indexů do modelu obdržíme koeficient determinace  $R^2 = 0,967$ , tzn. dostatečně vysoký. Při zahrnování dalších proměnných se koeficient determinace už o tolik nezvýší. Záleží nyní

na analytikovi, které proměnné do modelu zařadí a jak vysoký koeficient determinace bude požadovat.

Podíváme –li se na tabulku 1.1, která byla uvedena v kapitole 1, zjistíme, že společnosti podílející se na tvorbě indexu PX 50 k únoru 2006 reprezentují převážně obory jako je energetika, chemický, farmaceutický a gumárenský průmysl a peněžnictví, tzn. ty oborové indexy, které se podílely nejvíce na vysvětlení indexu PX 50 v letech 1996-1999.

Nyní sestrojíme model, kde vysvětlovanou proměnnou bude opět index PX 50 a vysvětlujícími proměnnými tři oborové indexy, které byly zrušeny k 28.2.2006. Budeme vycházet z toho, že není žádné časové zpoždění mezi indexem PX 50 a oborovými indexy, tzn. uvažovaná závislost je ve stejném časovém okamžiku. Tím máme opět regresi mezi indexem PX 50 a zbývajících třemi oborovými indexy. Data vezmeme od 5.1.2000 do 28.2.2006 a budeme pracovat opět s řadami prvních diferencí.

Model sestrojíme v programu STATISTICA. Výstup můžeme vidět níže.

#### *Výstup z programu STATISTICA*

Multiple Regression Results						
Dependent: Var5	Multiple R = ,77163390	F = 756,4498				
	R2= ,59541887	df = 3,1542				
No. of cases: 1546	adjusted R2= ,59463175	p = 0,000000				
	Standard error of estimate: 5,464701991					
Intercept: ,012314080	Std.Error: ,1397413	t( 1542) = ,08812	p = ,9298			
	BI07 beta= ,267	BI12 beta= ,597	BI16 beta= -,00			

*Tabulka 3.5*

	Beta	Std.Err.	B	Std.Err.	t(1542)	p-level
<b>Intercept</b>			0,012314	0,139741	0,08812	0,929792
<b>BI07</b>	0,266638	0,019220	0,486000	0,035031	13,87332	0,000000
<b>BI12</b>	0,596903	0,019186	0,278717	0,008959	31,11061	0,000000
<b>BI16</b>	-0,001667	0,016582	-0,002424	0,024110	-0,10054	0,919925

V tabulce 3.5 jsou obsaženy oborové indexy BI07 (chemický, farmaceutický a gumárenský průmysl, BI12 (energetiku) a BI16 (služby) včetně dílčích t-testů. Na základě těchto t-testů můžeme konstatovat, že proměnná BI16 je statisticky nevýznamná a proto do modelu nebude zahrnuta. Výsledný model má tedy tvar

$$PX_t = 0,267 \cdot BI07 + 0,597 \cdot BI12 \quad (3.1.3)$$

Koeficient determinace je velmi nízký, a to  $R^2 = 0,595$ . Můžeme tedy říci, že zbývající oborové indexy, index PX 50 moc dobře nevysvětlují. Jak již bylo výše řečeno, indexy BI07, BI12 a BI16 byly k 28.2.2006 naposledy propočteny.

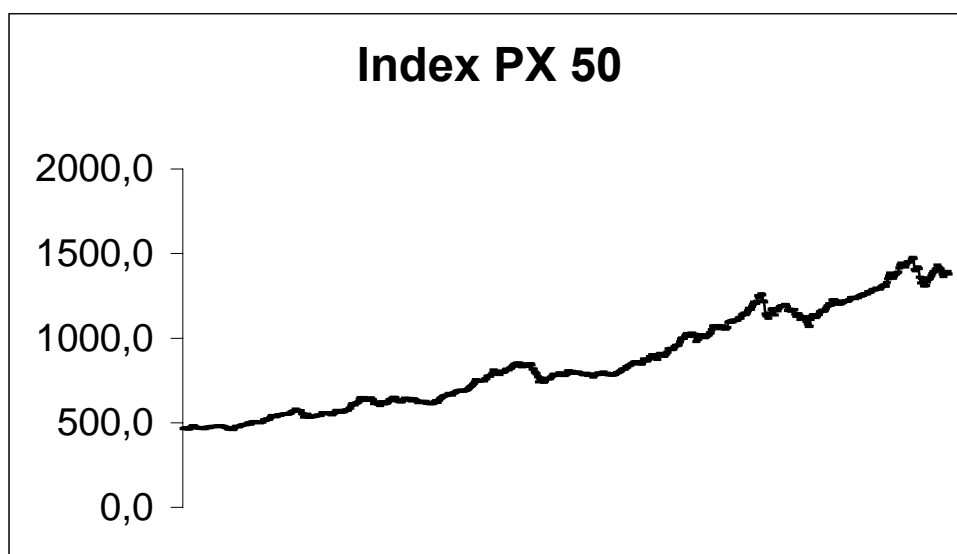
### 3.2. Index PX 50 versus index Dow Jones EURO STOXX 50

V kapitole 3.1. jsme se snažili zkonstruovat model, ve kterém bychom vysvětlili index PX 50 pomocí oborových indexů. V této části budeme zkoumat, zda index PX 50 není ovlivněn vývojem indexu Dow Jones EURO STOXX 50 (dále jen Euroindex, tento index by podrobně popsán v kapitole 1.1.). Tuto závislost se budeme snažit popsat opět pomocí modelu s přenosovou funkcí. Index PX 50 bude výstupní řadou a Euroindex bude vstupní řadou.

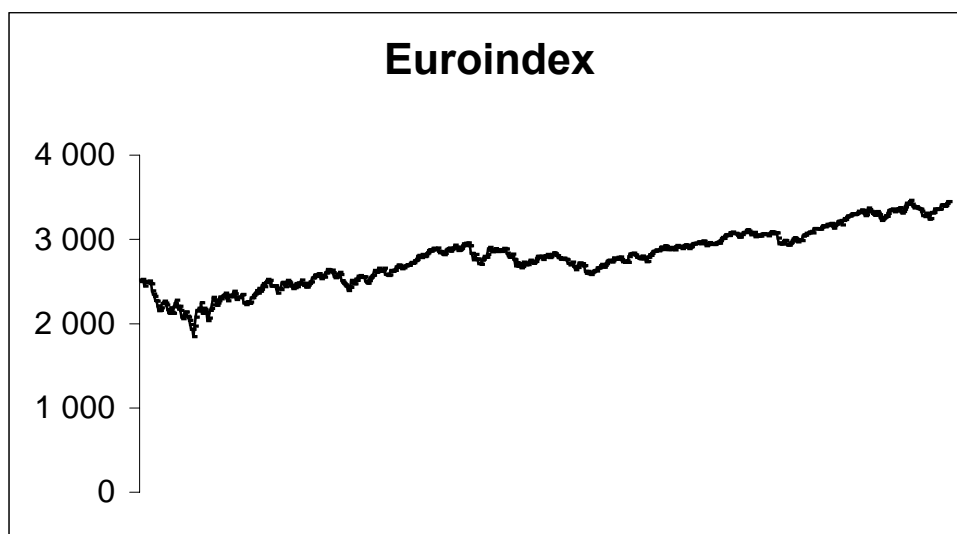
Za účelem analýzy vezmeme napozorovaná data od 2.1.2003 do 22.11.2005 u obou indexů. Jedná se o souvislé denní časové řady. U časové řady PX 50 denní data znamenají všechny dny, kdy BCPP obchodovala a u časové řady Euroindexu to znamená dny, kdy se obchodovalo s akcemi předních společností eurozóny. Zde se objevuje první problém, a to že každá řada je jiná. Za stejný časový úsek mají totiž řady indexu PX 50 a Euroindexu jiný počet pozorování, což je způsobeno tím, že obě burzy, jak pražská tak burza eurozóny, mají rozdílné dny, kdy se na nich obchoduje. Tato nesourodost je způsobena státními svátky, které nejsou v obou případech jednotné. Vzhledem k tomu, že teorie modelů s přenosovou funkcí předpokládá, že časové řady jsou pozorovány ve stejných časových bodech, je tento problém podstatný. Při jeho odstraňování budeme postupovat tak, že chybějící pozorování dopočítáme na základě lineární aproximace.

Nyní přistoupíme k vlastní analýze. Podíváme-li se na vývoj indexu PX 50 od 2.1.2003 do 22.11.2005 (máme tedy 744 pozorování), který je zaznamenán na obr.3.1 a na

vývoj indexu Euroindexu od 2.1.2003 do 22.11.2005 (opět máme 744 pozorování), který je zaznamenán na obr. 3.2, vidíme, že ani jedna řada není stacionární. K dosažení stacionarity použijeme diferencování. V tomto případě stačí pouze běžné diferencování a stačí obě řady diferencovat pouze jednou. Diferencování prvního řádu vychází z metody, při níž se posuzuje velikost odhadnutého rozptylu dané řady a rozptylu jejích diferencí. V našem případě řád difference  $d$  je roven jedné, neboť jednou diferencovaná řada dává nejmenší odhadnutý rozptyl. Pro stabilizaci rozptylu byla použita Cox-Boxova transformace.

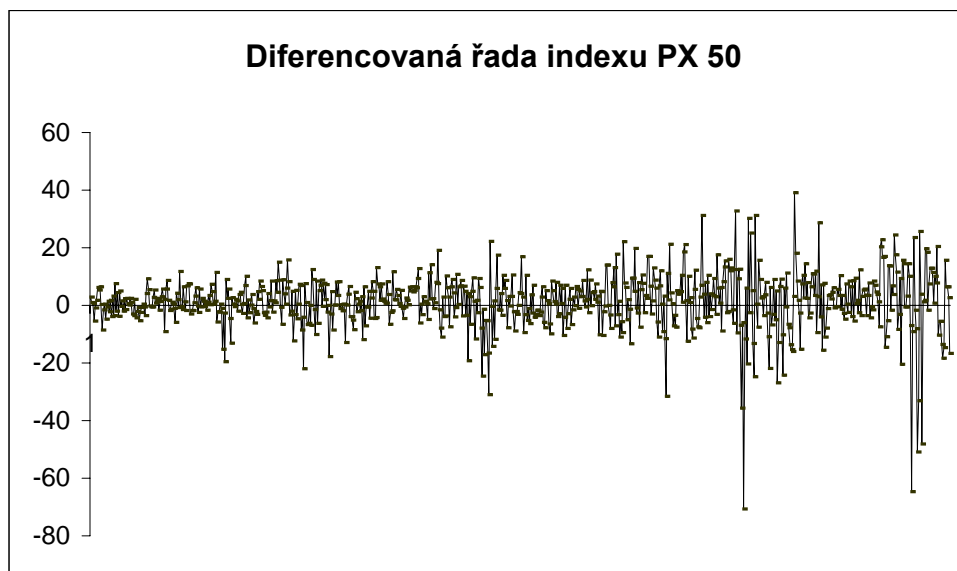


obr.3.1: Vývoj indexu PX 50

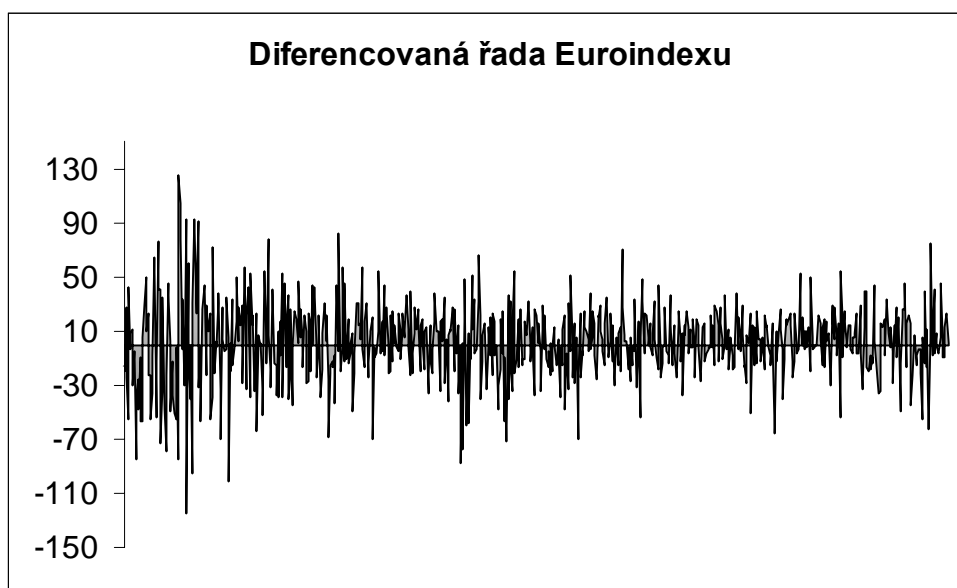


obr.3.2: Vývoj indexu Dow Jones Euro Stoxx 50

Diferencované řady jsou graficky znázorněny na obr.3.3 a obr.3.4 (obr.3.3 – diferencovaná řada indexu PX 50, obr.3.4 – diferencovaná řada Euroindexu).



*obr.3.3: Diferencovaná řada indexu PX 50*



*obr.3.4: Diferencovaná řada Euroindexu*

Nyní můžeme přejít ke konstrukci ARIMA modelu pro vstupní řadu (Euroindex) a výstupní řadu (index PX 50). Obě řady lze popsat modely AR(1). Výstupy z SCA můžeme vidět níže.

*Výstup z programu SCA: Model AR(1) pro řadu indexu PX 50*

SUMMARY FOR UNIVARIATE TIME SERIES MODEL -- PX									
-----									
VARIABLE	TYPE OF	ORIGINAL	DIFFERENCING						
	VARIABLE	OR CENTERED							
PX50	RANDOM	ORIGINAL	1 (1-B )						
-----									
PARAMETER	VARIABLE	NUM./	FACTOR	ORDER	CONS-	VALUE	STD	T	
LABEL	NAME	DENOM.			TRAIINT		ERROR	VALUE	
1	PHI1	PX50	AR	1	1	NONE	.1493	.0364	4.11
EFFECTIVE NUMBER OF OBSERVATIONS . .						742			
R-SQUARE . . . . .						.999			
RESIDUAL STANDARD ERROR. . . . .						.961154E+01			
--									

*Výstup z programu SCA: Model AR(1) pro řadu Euroindexu*

SUMMARY FOR UNIVARIATE TIME SERIES MODEL -- INDEX									
-----									
VARIABLE	TYPE OF	ORIGINAL	DIFFERENCING						
	VARIABLE	OR CENTERED							
EUROINDX	RANDOM	ORIGINAL	1 (1-B )						
-----									
PARAMETER	VARIABLE	NUM./	FACTOR	ORDER	CONS-	VALUE	STD	T	
LABEL	NAME	DENOM.			TRAIINT		ERROR	VALUE	
1	PHI1	EUROINDX	AR	1	1	NONE	-.0156	.0367	-.43
EFFECTIVE NUMBER OF OBSERVATIONS . .						742			
R-SQUARE . . . . .						.993			
RESIDUAL STANDARD ERROR. . . . .						.286067E+02			

Jak vidíme z výstupu programu SCA pro řadu indexu PX 50 obdržíme následující model AR(1)

$$(1 - B) \cdot (1 - 01493 \cdot B) \cdot Y_t = \varepsilon_t, \tag{3.2.1}$$

kde  $Y_t = index\ PX\ 50$ .

Koeficient determinace je roven 0,999 a reziduální směrodatná odchylka je  $\sigma_\varepsilon = 9,611$ .

Pro řadu Euroindexu dostáváme model AR(1) ve tvaru

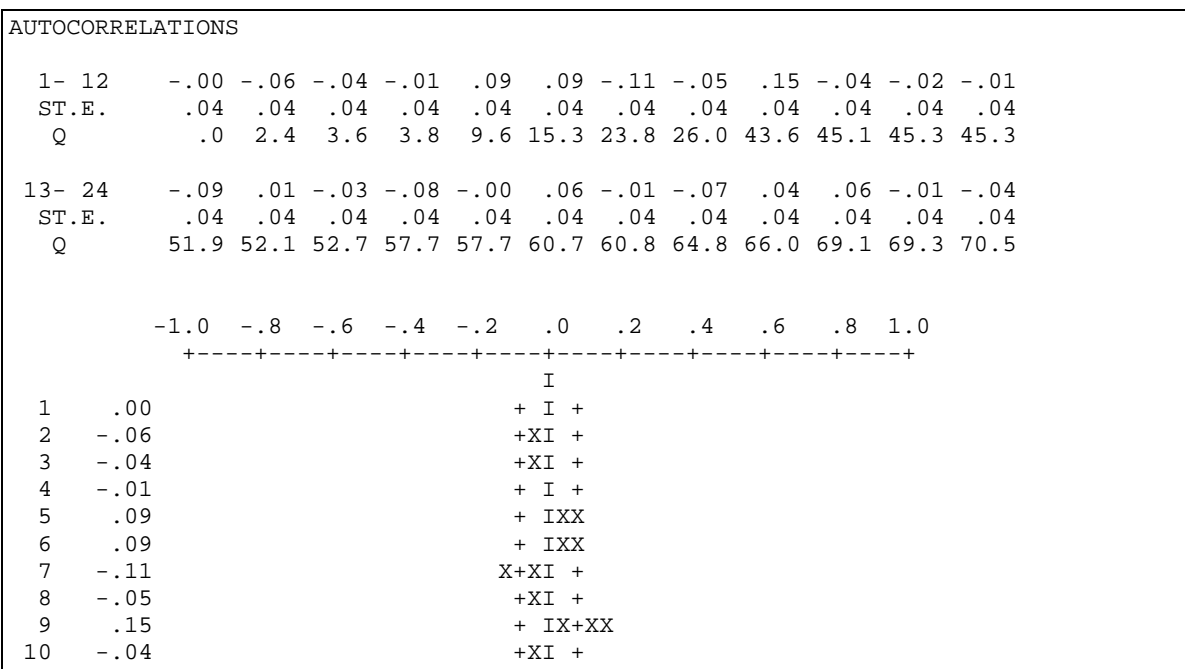
$$(1 - B) \cdot (1 + 0,0156 \cdot B) \cdot X_t = \varepsilon_t, \tag{3.2.2}$$

kde  $X_t = Euroindex$ .

Koeficient determinace je v tomto případě 0,993 a reziduální směrodatná odchylka je  $\sigma_\varepsilon = 28,6$ .

V obou případech byla ověřena adekvátnost modelu. Níže můžeme vidět na výstupu SCA, že u řady reziduí odpovídající indexu PX 50 dvě odhadnuté hodnoty autokorelační funkce překročily v absolutní hodnotě dvojnásobek příslušné směrodatné odchylky a u řady reziduí odpovídající Euroindexu k tomu došlo v jednom případě. Analogicky je to i u partiální autokorelační funkce.

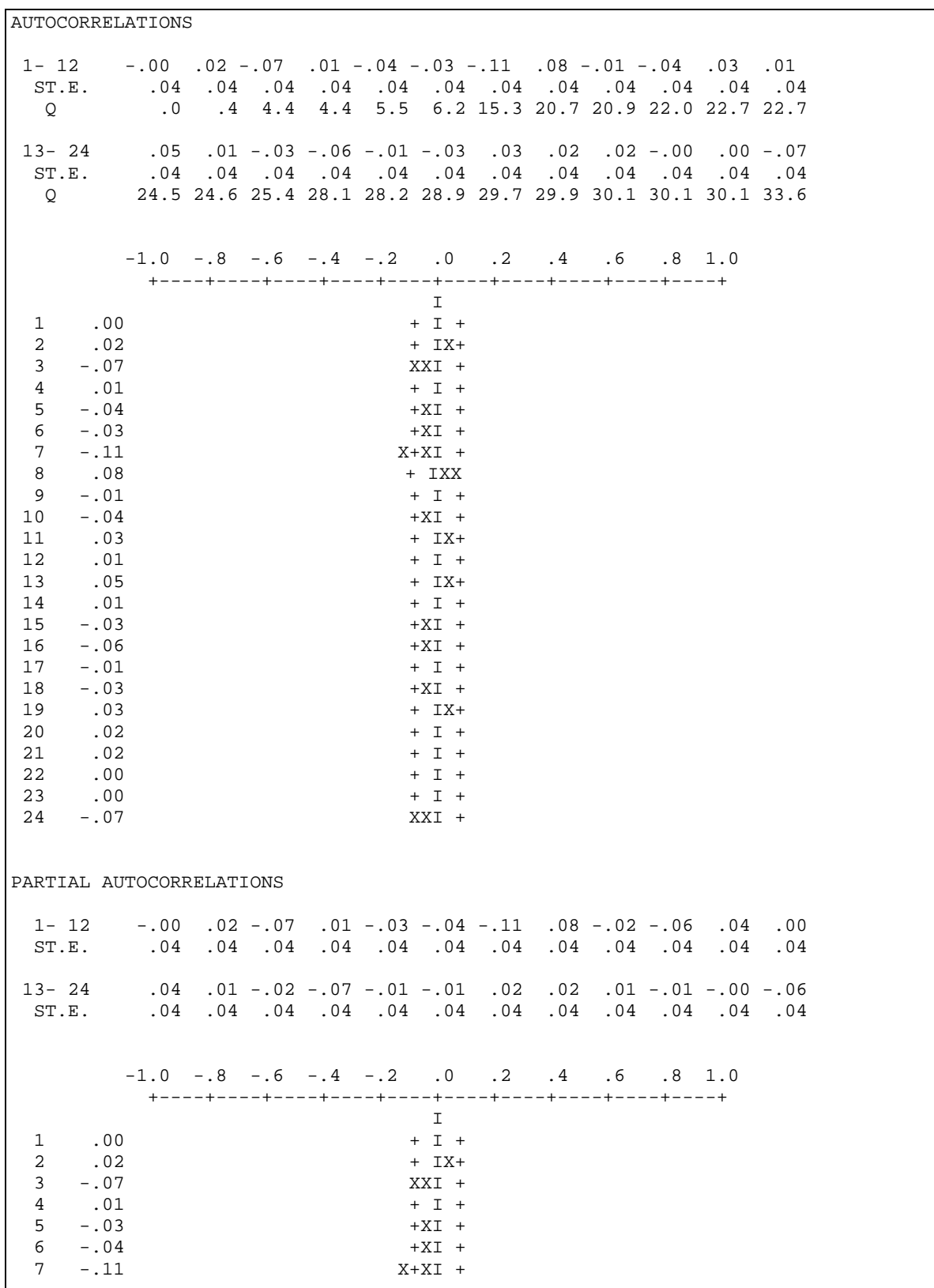
*Výstup z programu SCA: ACF a PACF reziduí řady PX 50*



11	-.02												+ I +
12	-.01												+ I +
13	-.09												XXI +
14	.01												+ I +
15	-.03												+XI +
16	-.08												XXI +
17	.00												+ I +
18	.06												+ IXX
19	-.01												+ I +
20	-.07												XXI +
21	.04												+ IX+
22	.06												+ IXX
23	-.01												+ I +
24	-.04												+XI +
PARTIAL AUTOCORRELATIONS													
1- 12	-.00	-.06	-.04	-.02	.08	.09	-.10	-.04	.16	-.06	-.03	.01	
ST.E.	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04
13- 24	-.07	-.02	-.06	-.05	-.00	.04	.01	-.08	.06	.10	-.06	-.04	
ST.E.	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04
	-1.0	-.8	-.6	-.4	-.2	.0	.2	.4	.6	.8	1.0		
1	.00												+ I +
2	-.06												+XI +
3	-.04												+XI +
4	-.02												+ I +
5	.08												+ IXX
6	.09												+ IXX
7	-.10												XXI +
8	-.04												+XI +
9	.16												+ IX+XX
10	-.06												XXI +
11	-.03												+XI +
12	.01												+ I +
13	-.07												XXI +
14	-.02												+ I +
15	-.06												XXI +
16	-.05												+XI +
17	.00												+ I +
18	.04												+ IX+
19	.01												+ I +
20	-.08												XXI +
21	.06												+ IX+
22	.10												+ IXX
23	-.06												+XI +
24	-.04												+XI +



Výstup z programu SCA: ACF a PACF reziduí řady Euroindexu



8	.08	+ IXX
9	-.02	+ I +
10	-.06	XXI +
11	.04	+ IX+
12	.00	+ I +
13	.04	+ IX+
14	.01	+ I +
15	-.02	+XI +
16	-.07	XXI +
17	-.01	+ I +
18	-.01	+ I +
19	.02	+ IX+
20	.02	+ IX+
21	.01	+ I +
22	-.01	+ I +
23	.00	+ I +
24	-.06	XXI +

Na základě testování ACF a PACF se tedy ukazuje, že zkonstruované modely nejsou dobré. Byly zkoušeny i další ARIMA modely, ale žádný z nich nevycházel uspokojivě. Musíme si uvědomit, že pracujeme s finančními časovými řadami, které mají určité charakteristické vlastnosti, které se promítly do modelu. Jednou z nich je právě nekonstantní podmíněný rozptyl. Je zřejmé, že lepší pro modelování budou modely ARCH a GARCH. Vzhledem k tomu, že zkoumáme vztah indexu PX 50 a Euroindexu, budeme v modelování pomocí přenosové funkce pokračovat, i když ARIMA modely naznačují, že tento model nebude příliš vhodný.

Nyní přejdeme ke konstrukci modelu s přenosovou funkcí. Budeme pracovat s následujícím tvarem modelu

$$Y_t = v_0 X_t + v_1 X_{t-1} + v_2 X_{t-2} + \dots + N_t, \quad (3.2.3)$$

kde

$Y_t$  (řada indexu PX 50) je výstupní řada,

$X_t$  (řada Euroindexu) je vstupní řada,

$N_t$  představuje poruchovou řadu.

Pro identifikaci vhodného modelu použijeme LTF metodu a sestrojíme předběžný model, čímž získáme výchozí odhady impulsních vah pro stanovení průběhu přenosové funkce. Z výstupu SCA vidíme, že významná je pouze váha  $v_0$ , ostatní váhy jsou nevýznamné. Tato skutečnost je zřejmá i z grafického záznamu vzájemné korelační funkce



CORRELATION BETWEEN EUROIND1 AND PX501 IS .26												
CROSS CORRELATION BETWEEN PX501(T) AND EUROIND1(T-L)												
1- 12	.06	.02	-.00	.03	.05	.00	-.02	-.04	-.04	-.00	.03	.02
ST.E.	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04
13- 24	-.06	-.01	-.02	-.01	.04	.04	.02	.04	-.01	.03	-.06	-.03
ST.E.	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04
CROSS CORRELATION BETWEEN EUROIND1(T) AND PX501(T-L)												
1- 12	.03	.00	-.03	.05	.02	.05	-.06	.00	.03	.02	-.04	-.00
ST.E.	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04
13- 24	-.04	-.01	.01	-.04	-.05	-.00	.01	.01	.01	.03	.00	.03
ST.E.	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04
-1.0 - .8 - .6 - .4 - .2 .0 .2 .4 .6 .8 1.0 +-----+-----+-----+-----+-----+-----+-----+-----+-----+												
I												
-24	.03											+ IX+
-23	.00											+ I +
-22	.03											+ IX+
-21	.01											+ I +
-20	.01											+ I +
-19	.01											+ I +
-18	.00											+ I +
-17	-.05											+XI +
-16	-.04											+XI +
-15	.01											+ I +
-14	-.01											+ I +
-13	-.04											+XI +
-12	.00											+ I +
-11	-.04											+XI +
-10	.02											+ IX+
-9	.03											+ IX+
-8	.00											+ I +
-7	-.06											+XI +
-6	.05											+ IX+
-5	.02											+ IX+
-4	.05											+ IX+
-3	-.03											+XI +
-2	.00											+ I +
-1	.03											+ IX+
0	.26											+ IX+XXXX
1	.06											+ IXX
2	.02											+ I +
3	.00											+ I +
4	.03											+ IX+
5	.05											+ IX+
6	.00											+ I +
7	-.02											+ I +
8	-.04											+XI +
9	-.04											+XI +
10	.00											+ I +
11	.03											+ IX+
12	.02											+ I +
13	-.06											+XI +
14	-.01											+ I +
15	-.02											+ I +
16	-.01											+ I +
17	.04											+ IX+
18	.04											+ IX+



Po dosazení odhadnutých parametrů  $\varphi_1, \varphi_9, \varphi_{20}$  a váhy  $v_0$  do (3.2.4) obdržíme konečný tvar modelu

$$(1 - 0,1456B - 0,1479B^9 + 0,0571B^{20}) \cdot (1 - B) \cdot Y_t = 0,0869 \cdot (1 - B) \cdot X_t + \varepsilon_t. \quad (3.2.5)$$

Adekvátnost modelu byla ověřována podle některých testů – vzájemně korelační funkcí reziduí pro přenosovou funkci a testem autokorelace pro poruchovou řadu. I když koeficient determinace  $R^2$  je poměrně vysoký, je roven 0,999, model (3.2.5) není příliš vhodný, což je zřejmé z předchozí analýzy. Vše naznačuje, že je třeba přejít k modelům volatility ARCH a GARCH.

Důvody, které vedou k tomu, abychom použili pro konstrukci modelu indexu PX 50 model ARCH, popř. GARCH jsou následující:

1. Modely ARIMA pro výstupní řadu PX 50 a vstupní řadu Euroindexu nejsou uspokojivé. U modelu AR(1) pro Euroindex vychází dokonce parametr  $\varphi_1$  dle t-testu statisticky nevýznamný.
2. Ze záznamu diferencovaných řad indexu PX 50 (obr.3.3) a Euroindexu (obr.3.4) je vidět, že v různých úsecích jednotlivých diferencovaných řad je různá variabilita.
3. V modelu (3.2.5) vidíme, že máme zpoždění 9 a 20. Můžeme tedy konstatovat, že časově vzdálené náhodné veličiny (náhodná veličina je v našem případě index PX 50) budou poměrně silně korelované. Jak již bylo poznamenáno v kapitole 1.2, tato vlastnost je typická v oblasti finančních časových řad. Máme tudíž model s dlouhou pamětí.
4. Hodnoty autokorelační funkce ACF a parciální autokorelační funkce PACF (0,8 či 0,9) jak pro poruchovou řadu tak pro rezidua leží na hranici intervalu spolehlivosti či těsně za ní.

Přes všechny uvedené výhrady k modelu s přenosovou funkcí musíme konstatovat, že sestrotit lepší model s přenosovou funkcí než model (3.2.5) nelze. Musíme si ale uvědomit, že nám přímo nešlo o sestrotění nejlepšího modelu pro řadu indexu PX 50, ale o modelování vztahu mezi indexem PX 50 a Euroindexem, což se nám podařilo.

Co se týká předpovědi na základě modelu (3.2.5), tak od ní ustoupíme, protože by nebyla s velkou pravděpodobností kvalitní. Důvodem nekvalitní předpovědi by byl sestrotěný neadekvátní model ARIMA pro vstupní řadu. Předpověď sestrotíme u modelu ARCH, popř. GARCH, na které celá analýza vede.

Při analýze modelu s přenosovou funkcí se při výpočtu vzájemné korelační funkce CCF ukázalo, že neexistuje zpoždění mezi řadami indexu PX 50 a Euroindexu, tzn. korelace je jen ve stejných časových bodech. Tím pádem se můžeme podívat, zda na modelování vztahu mezi těmito řadami nám nebude stačit regrese (podobně jako u modelování vztahu indexu PX 50 a oborových indexů).

V případě použití regrese obdržíme regresní model ve tvaru

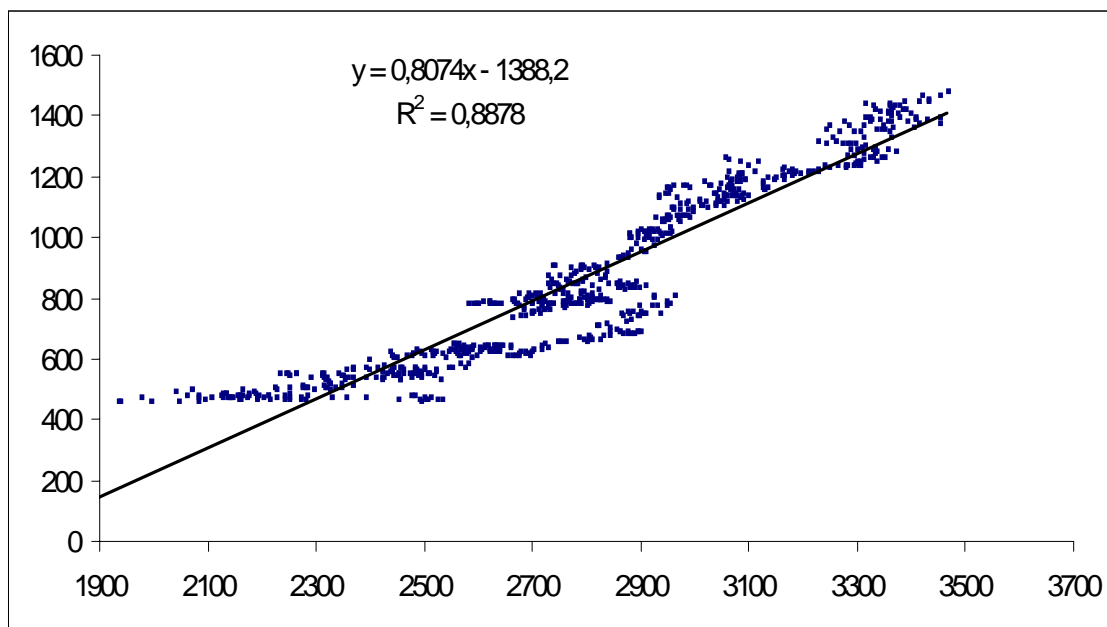
$$Y_t = 0,8074X_t - 1388,2, \quad (3.2.6)$$

kde  $Y_t$  = index PX 50 a  $X_t$  = Euroindex.

Koeficient determinace je  $R^2 = 0,8878$ . Regresní přímka je znázorněna na obr.3.5. Po zhlédnutí regresní přímky znázorněné na obr.3.5 můžeme konstatovat, že ani model (3.2.6) nebude vhodný. Pokusíme se tedy o určité vylepšení. Zvolíme dva přístupy. V prvním přístupu se budeme snažit data vyrovnat lomenou regresí. Z dat znázorněných na obr.3.5 je zřejmé, že v prvním úseku by se data dala vyrovnat jinou regresní přímkou než ve druhém úseku. Budeme požadovat, aby tyto přímky na sebe spojitě navázaly. K tomuto účelu vytvoříme následující regresní funkci proměnné  $X_t$

$$\begin{aligned} f(X_t) &= a + b \cdot X_t & t \leq j \\ &= a + b \cdot X_j + d \cdot (X_t - X_j) & t > j \end{aligned} \quad (3.2.7)$$

kde  $X_t = \text{Euroindex}$ ,  $a$ ,  $b$ ,  $d$  jsou regresní parametry (parametry  $b$ ,  $d$  jsou směrnice přímek vyrovnávací jednotlivé úseky) a bod  $j$  je bod zlomu, tzn. bod kde končí přímka se směrnici  $b$  a začíná přímka se směrnici  $d$ .



obr.3.5 – Regrese mezi indexem PX 50 a Euroindexem

Je zřejmé, že se problém redukuje na nalezení bodu zlomu  $j$ . Za bod zlomu vezmeme takové  $j$ , pro které platí

$$\hat{j} = \arg \min_{j=1, \dots, n} \{s_j^2\}, \quad (3.2.8)$$

kde

$$s_j^2 = \sum_{t=1}^n (Y_t - f(X_t))^2 \quad (3.2.9)$$

a  $Y_t = \text{index PX 50}$ ,  $n$  je počet pozorování, v našem případě  $n = 744$ .

Použijeme tedy metodu nejmenších čtverců na regresní funkci (3.2.7). Se získáním odhadu bodu zlomu  $\hat{j}$  získáme i odhady parametrů modelu (3.2.7)  $\hat{a}, \hat{b}, \hat{d}$ . Celou výše



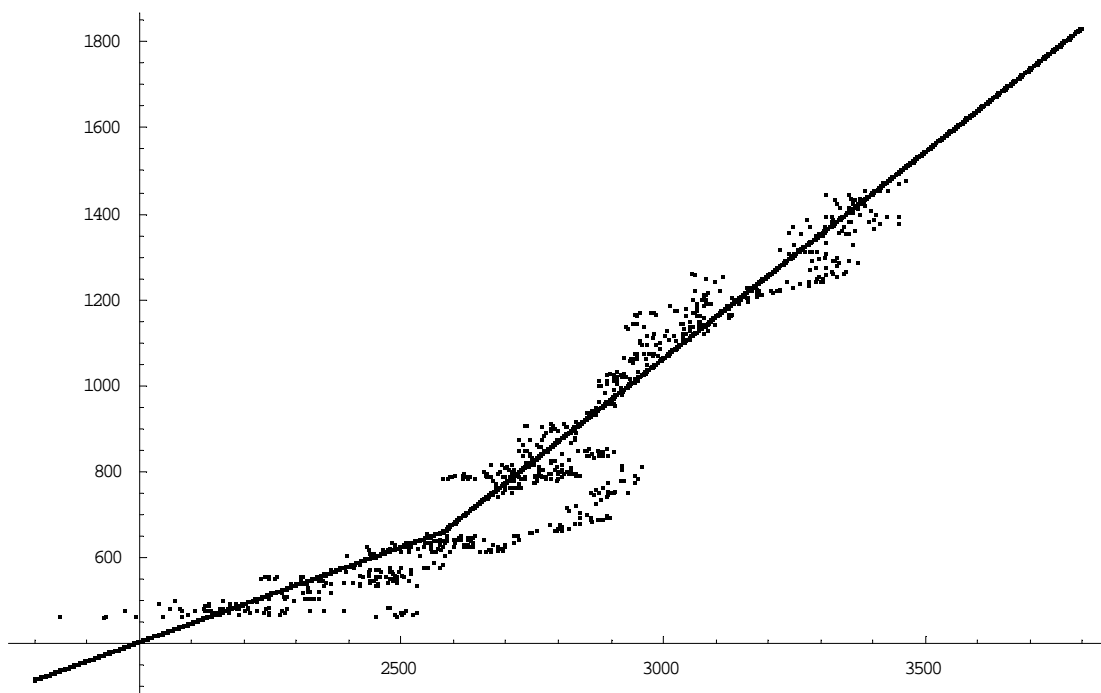
uvedenou proceduru provedeme v programu MATHEMATICA. Tuto proceduru musíme vytvořit, není v programu MATHEMATICA zabudovaná. Výstup najdeme v příloze.

Po provedení výše uvedené procedury obdržíme regresní model

$$\begin{aligned} Y_t &= 481,28 + 0,442 \cdot X_t & t \leq 415 \\ &= 481,28 + 0,442 \cdot X_j + 0,96 \cdot (X_t - X_j) & t > 415 \end{aligned} \quad (3.2.10)$$

kde  $Y_t$  = index PX 50 a  $X_t$  = Euroindex. Bod zlomu je  $j = 415$ .

Koeficient determinace se nám oproti modelu (3.2.6) podstatně zvýší,  $R^2 = 0,9589$ . Lomená regrese (3.2.10) je znázorněna na obr.3.6. Na tomto obrázku je vidět, že za bodem zlomu  $j=415$ , máme úsek, v němž se data odklání od obou regresních přímek.



obr.3.6: Lomená regrese

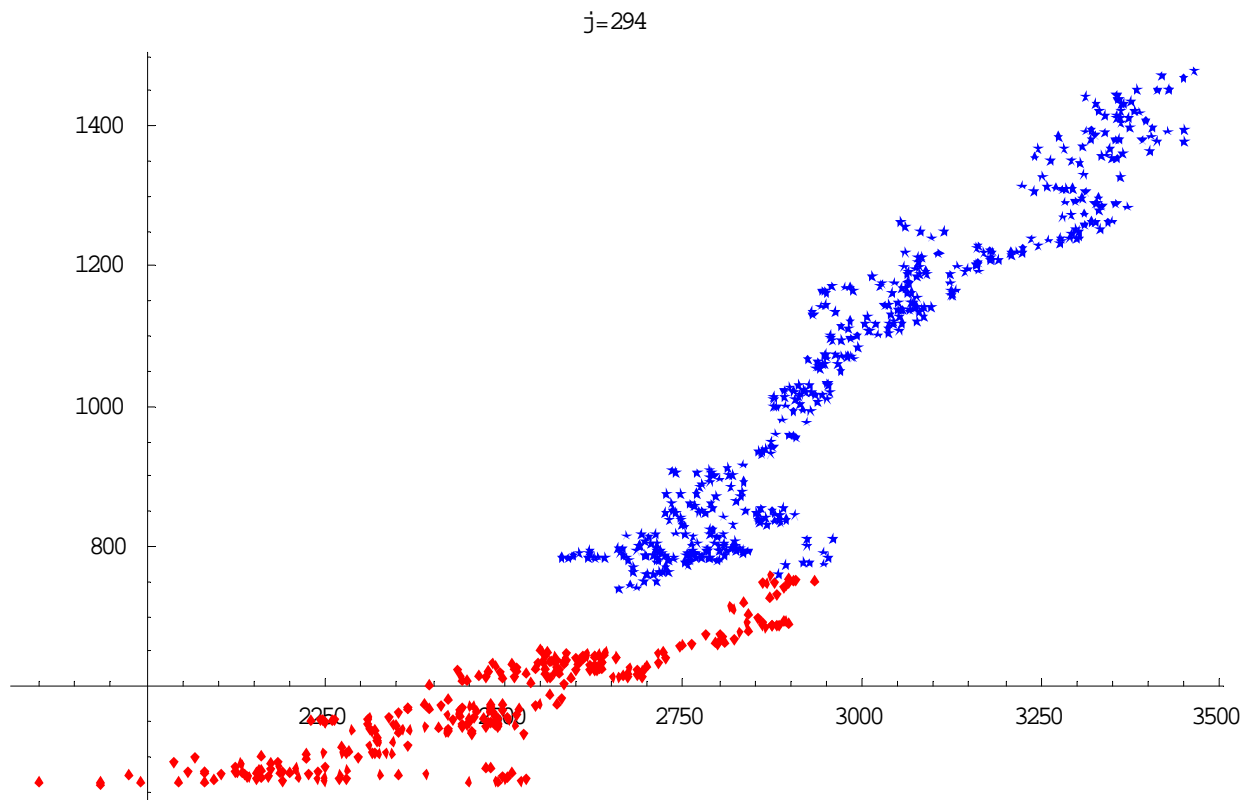
Pokusíme se tedy daná data vyrovnat dvěma regresními přímkami. Tyto přímkami nyní nebudou na sebe navazovat. Budeme hledat bod  $j$ , do kterého data vyrovnáme jednou regresní přímkou a od něho druhou regresní přímkou. Bod  $j$  budeme hledat empiricky. Postupně volíme  $j = 294, 304, 314, 324, 334, 344$ . Samozřejmě se zkoušela i jiná volba bodu  $j$ , ale nakonec se ukázaly výše uvedené body nejvhodnější. Na obr.3.7 až 3.12 můžeme vidět, že hledaný bod je  $j = 304$ . Body do okamžiku  $j$  jsou znázorněny červeně a po tomto okamžiku modře. Pokud  $j = 294$ , modré body zasahují do červené části. Pokud  $j = 314$  nebo  $324$  nebo  $334$  nebo  $344$  je tomu naopak, červené body zasahují do modré části. Pro  $j = 304$  máme červené a modré body od sebe separovány. V tomto časovém okamžiku došlo ke změně vývoje indexu z rostoucí tendence na klesající. Tudiž se můžeme domnívat, že v tomto časovém okamžiku došlo k nějaké významné události, která ovlivnila vývoj tohoto indexu.

V bodě  $j = 304$  tedy došlo k ovlivnění závislosti indexu. Z tohoto důvodu byla data v tomto bodě rozdělena na dvě části a pro každou část byl vytvořen regresní model. Regresní přímkami mají následující rovnice

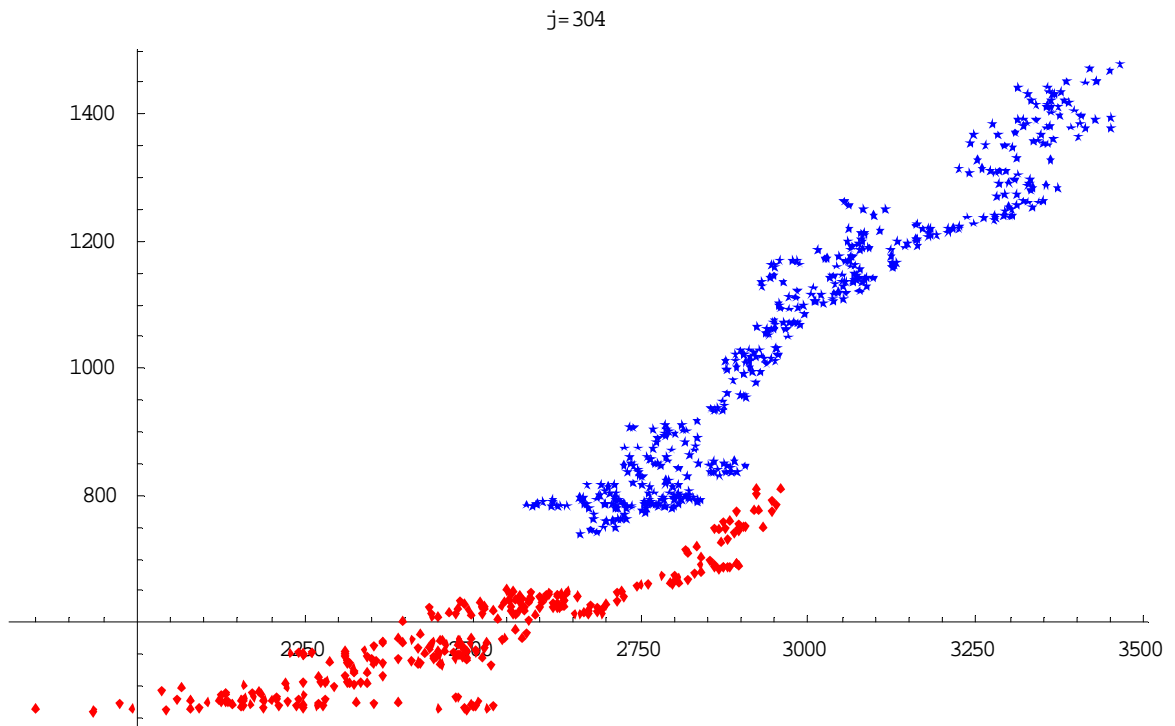
$$\begin{aligned} Y_t &= -225,858 + 0,325 \cdot X_t & t \leq 304 \\ Y_t &= -1644,3 + 0,902 \cdot X_t & t > 304 \end{aligned} \quad (3.2.11)$$

kde  $Y_t$  = index PX 50 a  $X_t$  = Euroindex. Koeficient determinace pro regresní model (3.2.11) se oproti modelu (3.2.10) opět zlepšil, platí  $R^2 = 0,969$ .

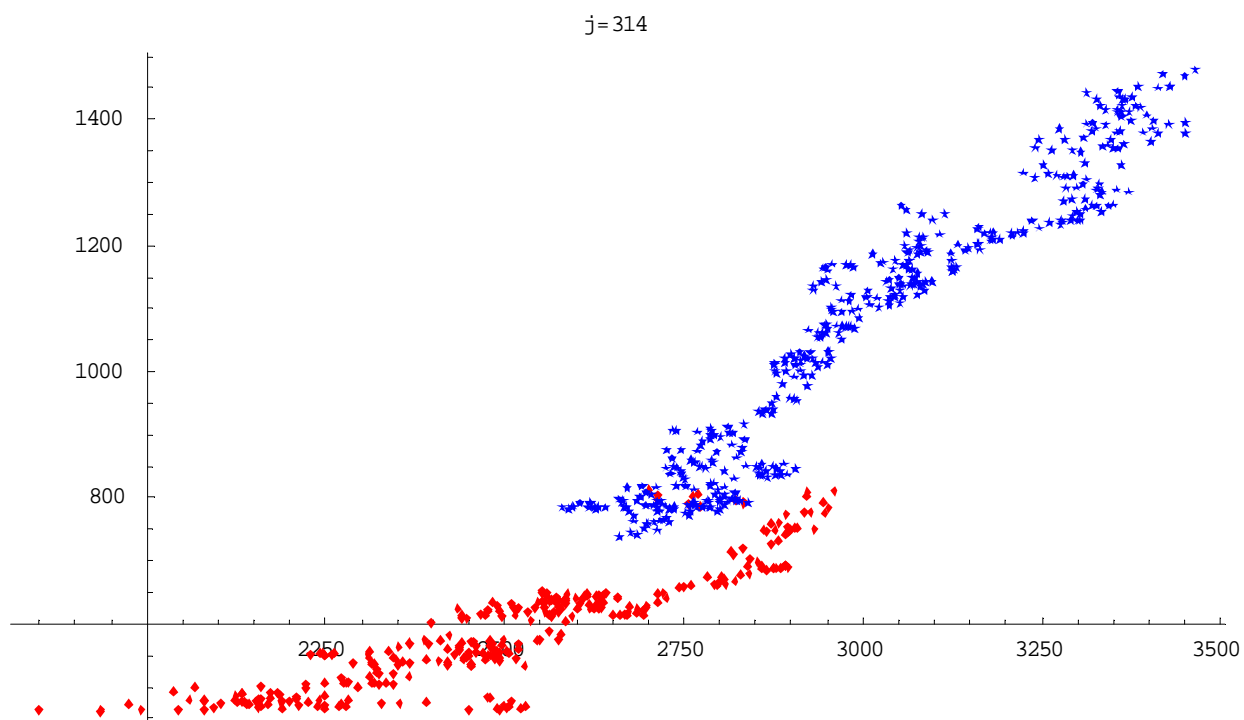
Z výše uvedeného je zřejmé, že i když uvažujeme regresi, kde vysvětlující proměnná je Euroindex a vysvětlovaná proměnná index PX 50, čas hraje při modelování této závislosti důležitou roli.



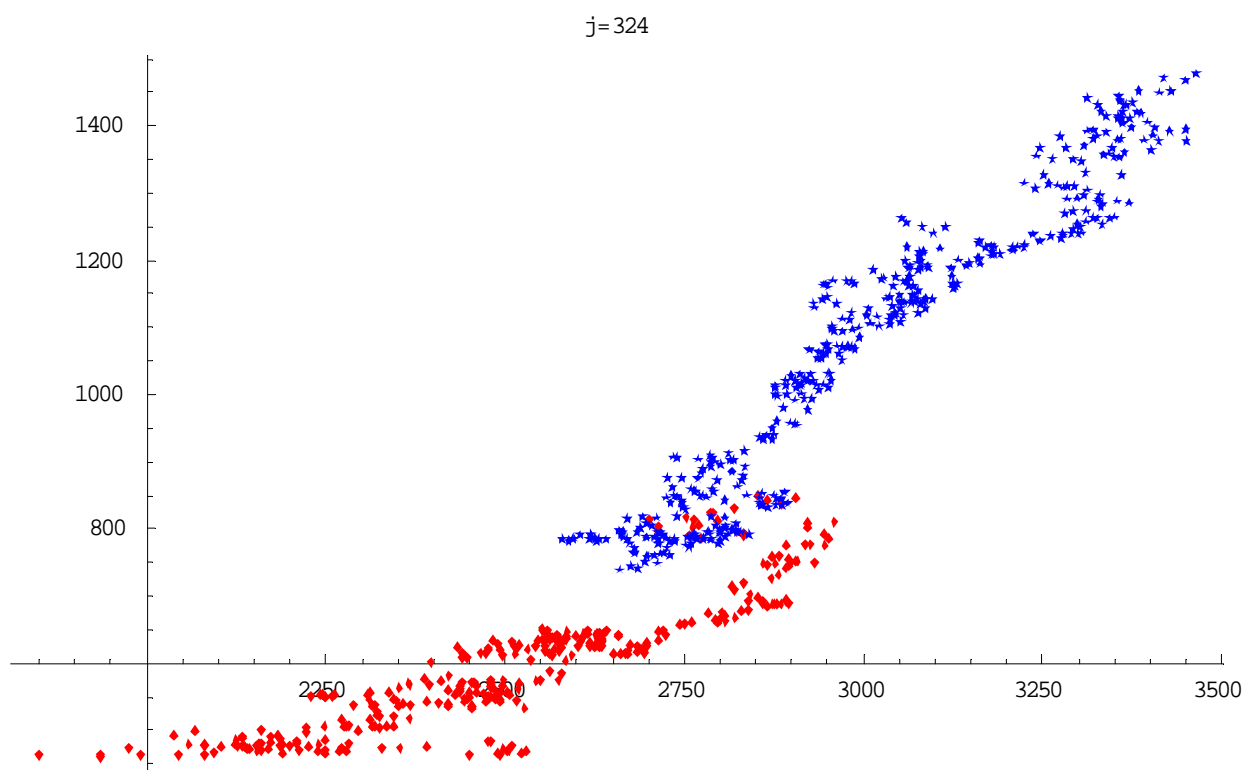
obr.3.7



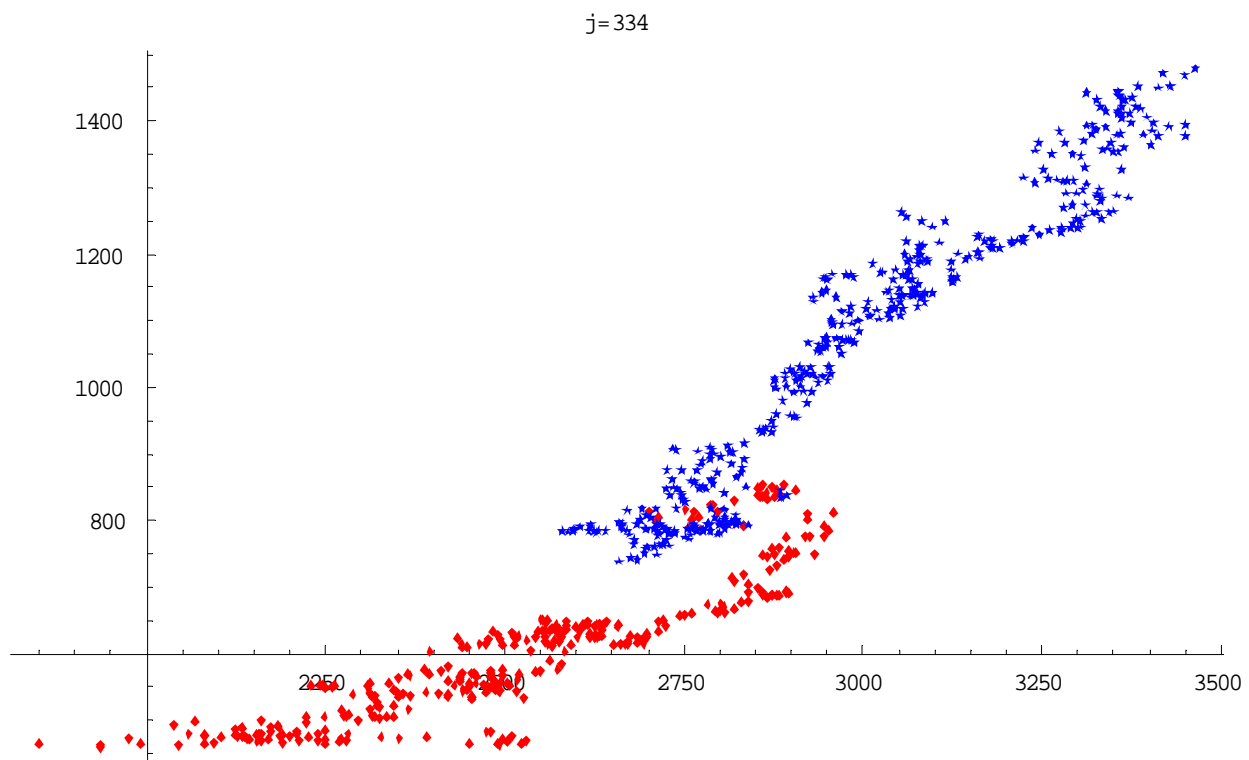
obr.3.8



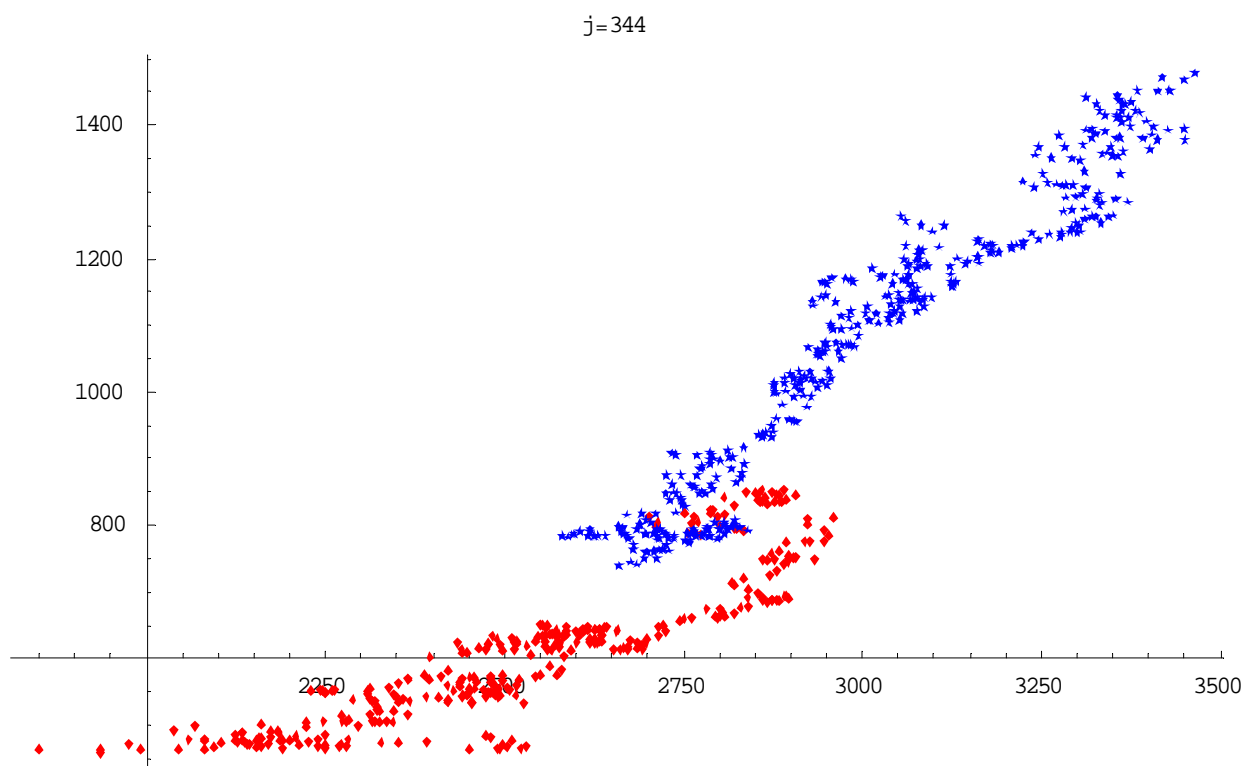
obr.3.9



obr.3.10



obr.3.11



obr.3.12

Na závěr ještě ukážeme, že trend obou regresních přímek je významně rozdílný. Tuto skutečnost vidíme z výstupu z programu MATHEMATICA, který je uveden v příloze. Za tímto účelem byla použita testovou statistiku

$$T = \frac{(b - b^*) \cdot \sqrt{n + n^* - 4}}{\sqrt{S_e + S_e^*} \cdot \sqrt{\frac{1}{\sum_1^n X_i^2 - n\bar{X}^2} + \frac{1}{\sum_1^{n^*} X_i^{*2} - n^*\bar{X}^{*2}}}}, \quad (3.2.12)$$

kteřá má Studentovo rozdělení o  $n + n^* - 4$  stupních volnosti, přičemž se vychází ze dvou nezávislých regresních modelů najednou, a to

$$\begin{aligned} Y_i &= \alpha + \beta X_i + e_i, & i = 1, 2, \dots, n \\ Y_i^* &= \alpha^* + \beta^* X_i^* + e_i^*, & i = 1, 2, \dots, n^* \end{aligned} \quad (3.2.13)$$

Hodnota testové statistiky  $T$  je v našem případě  $T = -35,53$ , kritická hodnota  $t_{n+n^*-4}^*(0,05) = 1,164$ , tudíž hypotézu o shodnosti trendu obou regresních přímek zamítáme.

Vraťme se ještě k otázce, proč došlo v daném okamžiku ke změně vývoje indexu z rostoucí tendence na klesající. Jak jsme již poznamenali, důvodem může být nějaká mimořádná událost. Touto událostí může být i pravidelný burzovní cyklus. V měsících březen - duben (do tohoto období spadá náš časový index  $j = 304$ ) dochází ke zveřejňování výročních zpráv všech podniků, tudíž i těch, které jsou zahrnuty do báze indexu PX 50 nebo Euroindexu. Zveřejnění výsledků hospodaření mohlo být mnohem horší než bylo očekávání, tím pádem mohlo dojít ke změně vývoje indexů z rostoucí tendence na klesající, neboť došlo zřejmě k poklesu tržních cen akcií zahrnutých do báze indexů. Vzhledem k tomu, že podniky dál fungují, dochází po určité době ke stabilizaci a k opětovné rostoucí tendenci vývoje indexů.

### 3.3. Modelování indexu PX 50 pomocí modelů volatility

V předchozích kapitolách jsme se snažili vytvořit modely s přenosovou funkcí, které měly vyjádřit závislost mezi finančními řadami indexu PX 50 a oborových indexů, popř. Euroindexu. Ukázalo se, že tyto modely nejsou příliš vhodné. Důvodem je proměnlivá volatility. Z obr.3.3 a obr.3.4 v kapitole 3.2., které znázorňují řady prvních diferencí indexu PX 50 a Euroindexu, vidíme, že v různých úsecích je volatility různá. Můžeme pozorovat výskyt segmentů s nízkou a naopak s vysokou volatilitou. Jak již bylo řečeno v úvodní části kapitoly 1.4, tyto změny volatility lze modelovat pomocí modelů volatility. V této kapitole se tedy budeme snažit najít takovýto model pro řadu indexu PX 50. Pro sestavení modelu použijeme software Eviews 5. Data pro statistické zpracování byla vzata od roku 1995 do dubna roku 2006. Budeme pracovat nikoliv s původní řadou indexu PX 50 (ani s řadou prvních diferencí), ale s mírou zisku uváděnou v %.

#### 3.3.1. Výběrové charakteristiky a rozdělení míry zisku indexu PX 50

Uvažujme míru zisku indexu PX 50 v obdobích znázorněných v tabulce 3.6. Začínáme rokem 1995, neboť teprve od konce roku 1994 se začalo na BCPP obchodovat pravidelně každý pracovní den v týdnu.

Tabulka 3.6 – Vývojové schéma časové řady míry zisku indexu PX 50 od 5.1.1995

<b>období</b>	<b>1-1</b>	<b>1-2</b>	<b>1-3</b>	<b>1-4</b>	<b>1-5</b>	<b>1-6</b>	<b>1-7</b>
<b>konečné datum</b>	15.12. 1995	31.12. 1996	30.12. 1997	30.12. 1998	30.12. 1999	29.12. 2000	28.12. 2001
<b>počet pozorování</b>	1-234	1-483	1-733	1-984	1-1238	1-1487	1737
<b>období</b>	<b>1-8</b>	<b>1-9</b>	<b>1-10</b>	<b>1-11</b>	<b>1-12</b>		
<b>konečné datum</b>	30.12. 2002	30.12. 2003	30.12. 2004	30.12. 2005	25.4. 2006		
<b>počet pozorování</b>	1-1987	1-2238	1-2485	1-2743	1-2824		

V tabulce 3.7 jsou uvedeny základní výběrové charakteristiky míry zisku indexu PX 50 za období uvedená v tabulce 3.6.

Tabulka 3.7 - Základní výběrové charakteristiky míry zisku indexu PX 50

<b>období</b>	<b>1-1</b>	<b>1-2</b>	<b>1-3</b>	<b>1-4</b>	<b>1-5</b>	<b>1-6</b>
<b>Základní charakteristiky</b>						
<i>průměr</i>	-0,11	0,00	-0,01	-0,03	0,00	0,00
<i>směrodatná odchylka</i>	0,79	0,76	0,85	1,06	1,12	1,2
<i>minimum</i>	-3,08	-4,14	-4,34	-6,83	-6,83	-6,83
<i>maximum</i>	2,49	2,83	4,41	4,85	5,99	5,99
<i>25%kvantil</i>	-0,52	-0,35	-0,41	-0,51	-0,53	-0,6
<i>75%kvantil</i>	0,31	0,36	0,42	0,49	0,52	0,58
<i>šikmost</i>	-0,2	-0,34	-0,32	-0,43	-0,2	-0,09
<i>špičatost</i>	2,57	7,09	3,62	4,08	3,8	2,59
<b>období</b>	<b>1-7</b>	<b>1-8</b>	<b>1-9</b>	<b>1-10</b>	<b>1-11</b>	<b>1-12</b>
<b>Základní charakteristiky</b>						
<i>průměr</i>	-0,01	0,00	0,01	0,03	0,04	0,04
<i>směrodatná odchylka</i>	1,24	1,26	1,23	1,21	1,2	1,19
<i>minimum</i>	-6,83	-6,83	-6,83	-6,83	-6,83	-6,83
<i>maximum</i>	5,99	5,99	5,99	5,99	5,99	5,99
<i>25%kvantil</i>	-0,67	-0,68	-0,65	-0,63	-0,59	-0,58
<i>75%kvantil</i>	0,60	0,64	0,66	0,68	0,69	0,69
<i>šikmost</i>	-0,07	-0,05	-0,09	-0,11	-0,18	-0,19
<i>špičatost</i>	2,21	1,85	1,9	1,97	2,16	2,20



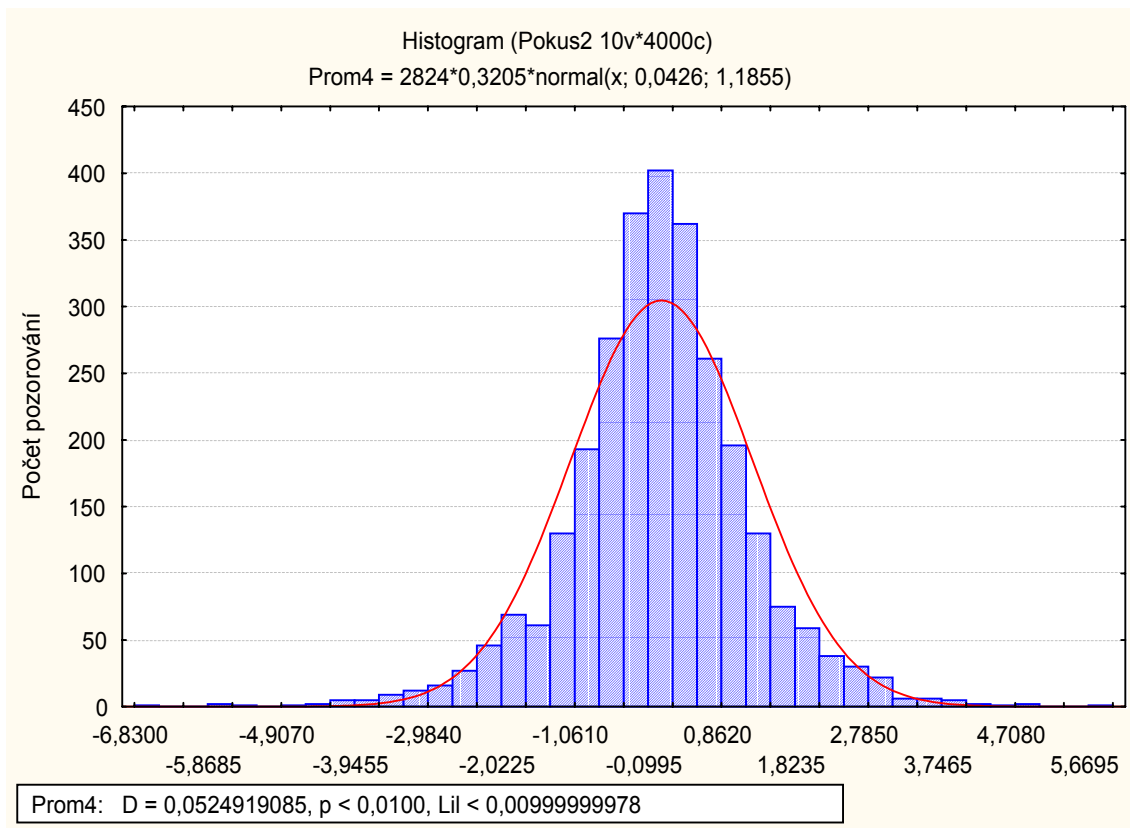
V tabulce 3.7 vidíme, jak se vyvíjí postupně základní charakteristiky míry zisku indexu PX 50. Mezi základní charakteristiky patří i šikmost a špičatost. Můžeme vidět, že odhad šikmosti je ve všech případech záporné číslo. Vzhledem k tomu, že odhad střední hodnoty je velmi malé číslo, lze konstatovat, že skutečné rozdělení je zešikmené tak, že záporné výnosy se objevují častěji než kladné výnosy. Co se týká odhadu špičatosti, je ve všech případech kladný, tudíž rozdělení míry zisku je špičatější než u normálního rozdělení. Na základě tohoto můžeme konstatovat, že nízké a záporné výnosy se objevují častěji než předpokládá normální rozdělení. Při výpočtu špičatosti vycházíme ze vztahu

$$Kurt = \frac{\mu_4}{\sigma^4} - 3, \quad (3.3.1)$$

kde  $\mu_4$  je čtvrtý centrální moment a  $\sigma^2$  je rozptyl.

Na obr.3.13 je znázorněna hustota normálního rozdělení a histogram získaný na základě našich dat. Ze znázornění je zřejmé, že naše data nepochází z normálního rozdělení. Toto jsme ověřili i pomocí testu normality  $\chi^2$  a Kolmogorova-Smirnovova testu (viz [5]). V obou případech byla hypotéza o normálním rozdělení zamítnuta.

Z výše uvedeného je tedy zřejmé, že rozdělení míry zisku je špičatější, zešikmené doleva a má „tlustší“ konce než rozdělení normální, tzn. že četnost výskytu extrémně vysokých kladných či záporných výnosů je vyšší než za předpokladu normality. Zároveň rozdělení míry zisku se vyznačuje „tenkým pásem“, kdy část hodnot uvažované náhodné veličiny leží s vysokou pravděpodobností blízko její střední hodnotě. V první kapitole jsme uvedli, že tato rozdělení se nazývají leptokurtickými rozděleními.



Obr.3.13 – Hustota normálního rozdělení a histogram pozorovaných dat

### 3.3.2. Model volatility

Nyní přistoupíme ke konstrukci modelu volatility pro míru zisku indexu PX 50. Postup konstrukce tohoto modelu je podrobně popsán v kapitole 1.4. Tam jsme také uvedli, že základním předpokladem modelů volatility je, že rozdělení náhodné veličiny  $e_t$  ze vztahu  $\varepsilon_t = e_t \cdot h_t^{1/2}$  za podmínky informace, která je k dispozici v čase  $t-1$ , je normované normální. Tím pádem je tedy rozdělení veličiny  $\varepsilon_t$  za podmínky informace, která je k dispozici v čase  $t-1$ , rovněž normální, avšak s podmíněným rozptylem, který se mění v závislosti na čase.

V praktických aplikacích se často zjišťuje, že špičatost standardizovaných reziduí je ale vyšší než 3 (což je špičatost normálního rozdělení), tudíž se někdy předpokládá, že veličina  $e_t$  má standardizované Studentovo rozdělení, které má špičatost větší než 3.

V této práci budeme předpokládat, že veličina  $e_t$  má právě standardizované Studentovo rozdělení. Pro toto rozdělení vychází model volatility řady míry zisku indexu PX 50 lépe. Tuto skutečnost můžeme vidět z níže uvedeného výstupu z programu Eviews 5. Hodnoty Akaikeho a Schwartzova kritéria jsou menší za předpokladu, že veličina  $e_t$  má standardizované Studentovo rozdělení.

*Výstup z programu Eviews 5 za předpokladu normality složky  $e_t$*

Dependent Variable: RETURN				
Method: ML - ARCH (Marquardt) - <b>Normal distribution</b>				
Date: 05/04/06 Time: 17:23				
Sample: 185 3007				
Included observations: 2823				
Convergence achieved after 21 iterations				
Variance backcast: ON				
GARCH = C(3) + C(4)*RESID(-1)^2 + C(5)*GARCH(-1)				
	Coefficient	Std. Error	z-Statistic	Prob.
RETURN(-1)	0.174704	0.020032	8.721342	0.0000
RETURN(-10)	0.053034	0.018580	2.854353	0.0043
Variance Equation				
C	0.025785	0.003986	6.468214	0.0000
RESID(-1)^2	0.114436	0.010385	11.01942	0.0000
GARCH(-1)	0.872240	0.011370	76.71717	0.0000
R-squared	0.012284	Mean dependent var		0.041821
Adjusted R-squared	0.010881	S.D. dependent var		1.184930
S.E. of regression	1.178465	Akaike info criterion		2.961043
Sum squared resid	3913.583	Schwarz criterion		2.971574
Log likelihood	-4174.513	Durbin-Watson stat		2.128077

Výstup z programu Eviews 5 za předpokladu, že  $e_t$  má standardizované Studentovo rozdělení

Dependent Variable: RETURN				
Method: ML - ARCH (Marquardt) - <b>Student's t distribution</b>				
Date: 05/04/06 Time: 17:24				
Sample: 185 3007				
Included observations: 2823				
Convergence achieved after 10 iterations				
Variance backcast: ON				
GARCH = C(3) + C(4)*RESID(-1)^2 + C(5)*GARCH(-1)				
	Coefficient	Std. Error	z-Statistic	Prob.
RETURN(-1)	0.169157	0.019084	8.863676	0.0000
RETURN(-10)	0.052336	0.017739	2.950325	0.0032
Variance Equation				
C	0.015296	0.004814	3.177280	0.0015
RESID(-1)^2	0.128079	0.015883	8.063955	0.0000
GARCH(-1)	0.872119	0.014060	62.02835	0.0000
T-DIST. DOF	6.749762	0.772187	8.741093	0.0000
R-squared	0.012887	Mean dependent var		0.041821
Adjusted R-squared	0.011135	S.D. dependent var		1.184930
S.E. of regression	1.178314	Akaike info criterion		2.913544
Sum squared resid	3911.191	Schwarz criterion		2.926180
Log likelihood	-4106.467	Durbin-Watson stat		2.116845

Na základě naší analýze dojdeme k tomu, že při konstrukci modelu indexu PX 50 kombinujeme autoregresní model řádu m-tého s modelem volatility GARCH(1,1), tj.

$$r_t = \varphi_1 \cdot r_{t-1} + \dots + \varphi_m \cdot r_{t-m} + \varepsilon_t, \quad \varepsilon_t = \sqrt{h_t} e_t \quad (3.3.2)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta \cdot h_{t-1} \quad e_t \text{ má standardizované}$$

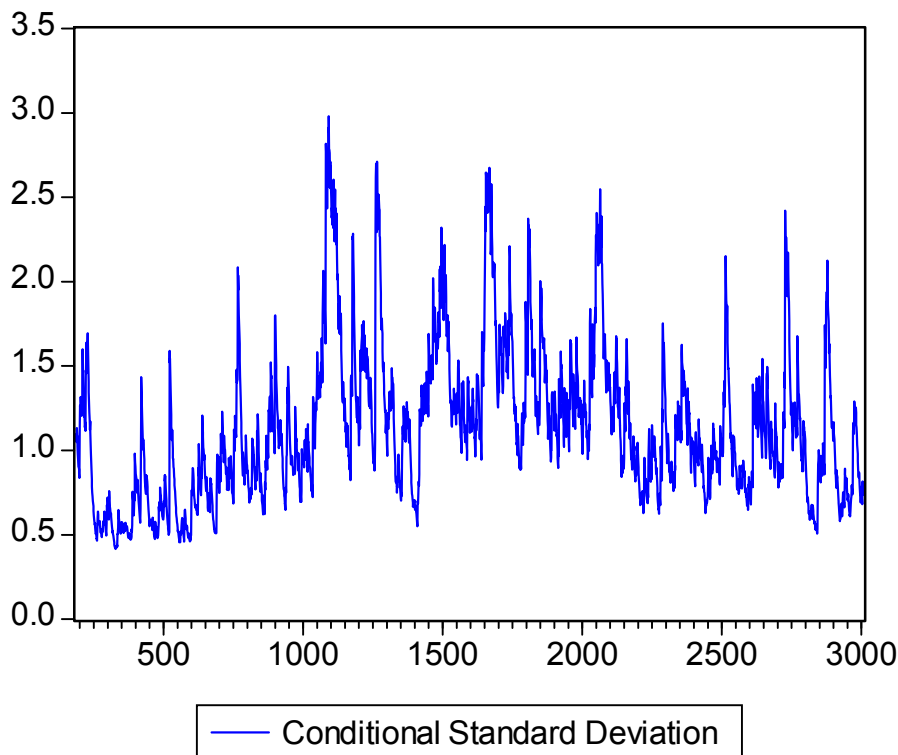
Studentovo rozdělení

a  $r_t$  je míra zisku indexu PX 50.

Z výstupu programu Eviews 5 vidíme, že výsledný model pro řadu míry zisku indexu PX 50 je  $AR(1,10) + GARCH(1,1)$ . Po dosazení parametrů  $\varphi_1, \varphi_{10}, \alpha_0, \alpha_1$  a  $\beta$  do (3.3.2) obdržíme

$$\begin{aligned} r_t &= 0,17 \cdot r_{t-1} + 0,052 \cdot r_{t-10} + \varepsilon_t \\ h_t &= 0,0153 + 0,128 \cdot \varepsilon_{t-1}^2 + 0,872 \cdot h_{t-1}. \end{aligned} \quad (3.3.3)$$

Na obr.3.14 je grafický záznam volatility procesu během let 1995-duben 2006. Vidíme, že dochází k segmentování. Vyskytují se segmenty s nízkou a naopak s vysokou volatilitou. Takové segmentování vzniká v důsledku toho, že předchozí vysoká (resp.nízká) volatilita vyvolává s velkou pravděpodobností také vysokou (resp.nízkou) volatilitu v následujícím čase.



Obr.3.14 – Podmíněná směrodatná odchylka během 1995- duben 2006

Model (3.3.3) je modelem míry zisku indexu PX 50 za období 1995 až duben 2006. Podívejme se nyní na model míry zisku indexu PX 50 za období uvedená v tabulce 3.8.

*Tabulka 3.8 – Vývojové schéma časové řady míry zisku indexu PX 50 od 5.1.1998 do 30.12.2003*

<b>období</b>	<b>1-1</b>	<b>1-2</b>	<b>1-3</b>	<b>1-4</b>	<b>1-5</b>	<b>1-6</b>
<b>konečné datum</b>	30.12. 1998	30.12. 1999	29.12. 2000	28.12. 2001	30.12. 2002	30.12. 2003
<b>počet pozorování</b>	1-251	1-493	1-741	1-991	1-1240	1-1490

Po celkové analýze opět dojdeme k modelu (3.3.2). Hodnoty významných parametrů za jednotlivá období jsou uvedena v tabulce 3.9 (viz [84]).

*Tabulka 3.9 – Odhadnuté statisticky významné parametry modelu (3.3.2) za období od 5.1.1998 do 30.12.2003*

<b>období</b>	<b>1-1</b>	<b>1-2</b>	<b>1-3</b>	<b>1-4</b>	<b>1-5</b>	<b>1-6</b>
<b>parametr</b>						
$\varphi_1$	0,18	0,16	0,14	0,11	0,08	0,08
$\varphi_{10}$	-	-	0,08	-	-	0,05
$\varphi_{31}$	-	-0,14	-0,09	-0,08	-0,07	-0,05
$\varphi_{33}$	-0,18	-	-	-	-	-
$\varphi_{42}$	-	-	-	0,09	0,06	0,05
$\alpha_0$	0,05	0,12	0,10	0,11	0,11	0,08
$\alpha_1$	0,09	0,18	0,13	0,11	0,10	0,10
$\beta$	0,89	0,77	0,83	0,84	0,85	0,86

Po shlédnutí tabulky 3.9 můžeme konstatovat, že v modelu 3.3.2 za daná období máme zpoždění až 42, tzn. časově vzdálené náhodné veličiny (náhodná veličina je v našem případě míra zisku indexu PX 50) jsou poměrně silně korelované.

### **3.4. Předpověď volatility pro míru zisku indexu PX 50, indexu Standard & Poor's 500 a indexu Dow Jones EURO STOXX 50**

Důležitou součástí některých metod souvisejících s regulací finančních rizik je předpověď budoucí volatility různých měr zisku. Předpověď budoucího růstu volatility určitého finančního nástroje může být pro investora signálem, aby se této složky investičního či tržního portfolia rychle zbavil a tak redukoval svou budoucí expozici vůči příslušnému riziku. Předpověď budoucí volatility je také důležitá v případech, když cena finančního nástroje závisí na volatilitě předpověditelným způsobem (např. u opcí).

V této části budeme konstruovat předpověď volatility pro míru zisku indexu PX 50 jednak o jeden krok dopředu a jednak o  $T$  kroků dopředu. V prvním případě bude předpověď konstruovaná pro čas  $t+1$  v čase  $t$  a předpověď volatility budeme značit  $\sigma_{t+1|t}$ . Ve druhém případě bude předpověď konstruovaná pro čas  $t+T$  v čase  $t$  a budeme jí značit  $\sigma_{t+T|t}$ . V obou případech je předpověď založena na pozorovaných hodnotách míry zisku  $r_t, r_{t-1}, \dots$ . Při výpočtu se budeme řídit vzorci, které jsme uvedli v kapitoly 2.2.

Nyní se podíváme na jednodenní předpovědi volatility míry zisku indexu PX 50 a míry zisku indexu Standard & Poor's 500 po teroristickém útoku 11.9.2001 na World Trade Center v New Yorku. Tyto jednodenní předpovědi volatility jsou uvedeny v tabulkách 3.10 a 3.11 v rozmezí od 11.9.2001 do 18.10.2001. Jednodenní předpovědi volatility míry zisku u obou indexů až do konce roku jsou uvedeny v tabulkách 1 a 2 v příloze. Z tabulky 3.10 je vidět, že jednodenní předpovědi volatility míry zisku indexu PX 50 během uvedeného období postupně rostly, z tabulky 1 v příloze můžeme vidět, že po 18.10.2001 tyto předpovědi postupně začaly klesat. Analogicky se chovají jednodenní předpovědi volatility míry zisku indexu Standard & Poor's 500 (viz tabulka 3.11 a tabulka 2 v příloze).

Tabulka 3.10 – Předpovědi volatility pomocí exponenciálního vyrovnávání pro míru zisku indexu PX 50 od 11.9.2001 do 18.10.2001

Datum	t	Index PX 50	$r_t(\%)$	$r_t^2$	$\sigma_{t+1 t}^2$	$\sigma_{t+1 t}$
11.9.2001	0	340,0	-1,762	3,106	3,106	1,762
12.9.2001	1	330,5	-2,794	7,807	3,388	1,841
13.9.2001	2	335,4	1,483	2,198	3,317	1,821
14.9.2001	3	322,9	-3,727	13,890	3,951	1,988
17.9.2001	4	320,1	-0,867	0,752	3,759	1,939
18.9.2001	5	321,2	0,344	0,118	3,541	1,882
19.9.2001	6	330,9	3,020	9,120	3,876	1,969
20.9.2001	7	327,8	-0,937	0,878	3,696	1,922
21.9.2001	8	325,2	-0,793	0,629	3,512	1,874
24.9.2001	9	328,0	0,861	0,741	3,346	1,829
25.9.2001	10	332,7	1,433	2,053	3,268	1,808
26.9.2001	11	337,0	1,292	1,670	3,172	1,781
27.9.2001	12	331,9	-1,513	2,290	3,119	1,766
1.10.2001	13	333,9	0,603	0,363	2,954	1,719
2.10.2001	14	339,6	1,707	2,914	2,952	1,718
3.10.2001	15	334,7	-1,443	2,082	2,899	1,703
4.10.2001	16	340,4	1,703	2,900	2,899	1,703
5.10.2001	17	340,3	-0,029	0,001	2,725	1,651
8.10.2001	18	343,1	0,823	0,677	2,603	1,613
9.10.2001	19	344,2	0,321	0,103	2,453	1,566
10.10.2001	20	350,6	1,859	3,457	2,513	1,585
11.10.2001	21	359,2	2,453	6,017	2,723	1,650
12.10.2001	22	354,6	-1,281	1,640	2,658	1,630
15.10.2001	23	351,3	-0,931	0,866	2,551	1,597
16.10.2001	24	352,6	0,370	0,137	2,406	1,551
17.10.2001	25	359,1	1,843	3,398	2,465	1,570
18.10.2001	26	359,4	0,084	0,007	2,318	1,522



Tabulka 3.11 – Předpovědi volatility pomocí exponenciálního vyrovnávání pro míru zisku indexu Standard & Poor's 500 od 10.9.2001 do 18.10.2001

Datum	t	Index Standard & Poor's 500	$r_t(\%)$	$r_t^2$	$\sigma_{t+1 t}^2$	$\sigma_{t+1 t}$
10.9.2001	0	1 092,54	0,623	0,388	0,388	0,623
17.9.2001	1	1 038,77	-4,922	24,222	1,818	1,348
18.9.2001	2	1 032,74	-0,580	0,337	1,729	1,315
19.9.2001	3	1 016,10	-1,611	2,596	1,781	1,334
20.9.2001	4	984,54	-3,106	9,647	2,253	1,501
21.9.2001	5	965,80	-1,903	3,623	2,335	1,528
24.9.2001	6	1 003,45	3,898	15,197	3,107	1,763
25.9.2001	7	1 012,27	0,879	0,773	2,967	1,722
26.9.2001	8	1 007,04	-0,517	0,267	2,805	1,675
27.9.2001	9	1 018,61	1,149	1,320	2,716	1,648
28.9.2001	10	1 040,94	2,192	4,806	2,841	1,686
1.10.2001	11	1 038,55	-0,230	0,053	2,674	1,635
2.10.2001	12	1 051,33	1,231	1,514	2,604	1,614
3.10.2001	13	1 072,28	1,993	3,971	2,686	1,639
4.10.2001	14	1 069,62	-0,248	0,062	2,529	1,590
5.10.2001	15	1 071,38	0,165	0,027	2,379	1,542
8.10.2001	16	1 062,44	-0,834	0,696	2,278	1,509
9.10.2001	17	1 056,75	-0,536	0,287	2,158	1,469
10.10.2001	18	1 080,99	2,294	5,262	2,344	1,531
11.10.2001	19	1 097,43	1,521	2,313	2,343	1,531
12.10.2001	20	1 091,65	-0,527	0,277	2,219	1,490
15.10.2001	21	1 089,98	-0,153	0,023	2,087	1,445
16.10.2001	22	1 097,54	0,694	0,481	1,991	1,411
17.10.2001	23	1 077,09	-1,863	3,472	2,079	1,442
18.10.2001	24	1 068,61	-0,787	0,620	1,992	1,411

Vraťme se do současnosti. V tabulce 3.12, 3.13, 3.14 jsou vypočteny jednodenní předpovědi volatility míry zisku indexu PX 50, indexu Standard & Poor's 500 a indexu Dow Jones Euro Stoxx 50. Tyto předpovědi jsou napočítány od 17.3.2006 až do současnosti, tj. 25.4.2006. Vidíme, že jednodenní předpovědi volatility míry zisku se u všech indexů chovají stejně, tzn. rostou.

*Tabulka 3.12 - Předpovědi volatility pomocí exponenciálního vyrovnávání pro míru zisku indexu PX 50 od 17.3.2006 do 25.4.2006*

Datum	t	Index PX 50	$r_t(\%)$	$r_t^2$	$\sigma_{t+1 t}^2$	$\sigma_{t+1 t}$
17.3.2006	0	1539,7	0,085	0,007	0,007	0,085
20.3.2006	1	1554,6	0,968	0,936	0,063	0,251
21.3.2006	2	1544,3	-0,663	0,439	0,085	0,292
22.3.2006	3	1539,9	-0,285	0,081	0,085	0,292
23.3.2006	4	1544,9	0,325	0,105	0,086	0,294
24.3.2006	5	1538,5	-0,414	0,172	0,092	0,303
27.3.2006	6	1535,7	-0,182	0,033	0,088	0,297
28.3.2006	7	1527,1	-0,560	0,314	0,102	0,319
29.3.2006	8	1516,4	-0,701	0,491	0,125	0,353
30.3.2006	9	1530,5	0,930	0,865	0,169	0,411
31.3.2006	10	1523,9	-0,431	0,186	0,170	0,413
3.4.2006	11	1532,2	0,545	0,297	0,178	0,422
4.4.2006	12	1534,5	0,150	0,023	0,169	0,411
5.4.2006	13	1538,7	0,274	0,075	0,163	0,404
6.4.2006	14	1546,4	0,500	0,250	0,168	0,410
7.4.2006	15	1549,6	0,207	0,043	0,161	0,401
10.4.2006	16	1536,8	-0,826	0,682	0,192	0,438
11.4.2006	17	1518,1	-1,217	1,481	0,269	0,519
12.4.2006	18	1516,6	-0,099	0,010	0,254	0,504
13.4.2006	19	1509,6	-0,462	0,213	0,251	0,501
14.4.2006	20	1515,3	0,378	0,143	0,245	0,495
18.4.2006	21	1513,0	-0,152	0,023	0,231	0,481
19.4.2006	22	1533,8	1,375	1,890	0,331	0,575
20.4.2006	23	1547,9	0,919	0,845	0,362	0,602
21.4.2006	24	1550,9	0,194	0,038	0,342	0,585
24.4.2006	25	1546,1	-0,309	0,096	0,328	0,572
25.4.2006	26	1543,3	-0,181	0,033	0,310	0,557

Tabulka 3.13 - Předpovědi volatility pomocí exponenciálního vyrovnávání pro míru zisku indexu Standard & Poor's 500 od 17.3.2006 do 25.4.2006

Datum	t	Index Standard & Poor's 500	$r_t(\%)$	$r_t^2$	$\sigma_{t+1 t}^2$	$\sigma_{t+1 t}$
17.3.2006	0	1 307,25	0,147	0,022	0,022	0,147
20.3.2006	1	1 305,08	-0,166	0,028	0,022	0,148
21.3.2006	2	1 297,23	-0,601	0,362	0,042	0,206
22.3.2006	3	1 305,04	0,602	0,362	0,062	0,248
23.3.2006	4	1 301,67	-0,258	0,067	0,062	0,249
24.3.2006	5	1 302,95	0,098	0,010	0,059	0,242
27.3.2006	6	1 301,61	-0,103	0,011	0,056	0,236
28.3.2006	7	1 293,23	-0,644	0,415	0,077	0,278
29.3.2006	8	1 302,89	0,747	0,558	0,106	0,326
30.3.2006	9	1 300,25	-0,203	0,041	0,102	0,320
31.3.2006	10	1 294,83	-0,417	0,174	0,107	0,326
3.4.2006	11	1 297,81	0,230	0,053	0,103	0,322
4.4.2006	12	1 305,93	0,626	0,391	0,121	0,347
5.4.2006	13	1 311,56	0,431	0,186	0,125	0,353
6.4.2006	14	1 309,04	-0,192	0,037	0,119	0,345
7.4.2006	15	1 295,50	-1,034	1,070	0,176	0,420
10.4.2006	16	1 296,60	0,085	0,007	0,166	0,408
11.4.2006	17	1 286,57	-0,774	0,598	0,192	0,438
12.4.2006	18	1 288,12	0,120	0,015	0,181	0,426
13.4.2006	19	1 289,12	0,078	0,006	0,171	0,413
17.4.2006	20	1 285,33	-0,294	0,086	0,166	0,407
18.4.2006	21	1 307,65	1,737	3,015	0,337	0,580
19.4.2006	22	1 309,93	0,174	0,030	0,318	0,564
20.4.2006	23	1 311,46	0,117	0,014	0,300	0,548
21.4.2006	24	1 311,28	-0,014	0,000	0,282	0,531
24.4.2006	25	1 308,11	-0,242	0,058	0,269	0,518
25.4.2006	26	1 301,74	-0,487	0,237	0,267	0,517

Tabulka 3.14 - Předpovědi volatility pomocí exponenciálního vyrovnávání pro míru zisku indexu Dow Jones Euro Stoxx 50 od 17.3.2006 do 25.4.2006

Datum	t	Index Dow Jones Euro Stoxx 50	$r_t(\%)$	$r_t^2$	$\sigma_{t+1 t}^2$	$\sigma_{t+1 t}$
17.3.2006	0	3 832,43	-0,253	0,064	0,064	0,253
20.3.2006	1	3 842,10	0,252	0,064	0,065	0,256
21.3.2006	2	3 848,17	0,158	0,025	0,064	0,254
27.3.2006	3	3 827,93	-0,526	0,277	0,078	0,280
28.3.2006	4	3 811,22	-0,437	0,191	0,087	0,294
29.3.2006	5	3 826,30	0,396	0,157	0,093	0,304
30.3.2006	6	3 874,61	1,263	1,594	0,184	0,430
31.3.2006	7	3 853,74	-0,539	0,290	0,195	0,441
3.4.2006	8	3 878,64	0,646	0,417	0,212	0,460
4.4.2006	9	3 850,11	-0,736	0,541	0,236	0,486
5.4.2006	10	3 863,92	0,359	0,129	0,234	0,484
6.4.2006	11	3 861,29	-0,068	0,005	0,225	0,474
7.4.2006	12	3 823,11	-0,989	0,978	0,275	0,524
10.4.2006	13	3 843,52	0,534	0,285	0,281	0,530
11.4.2006	14	3 788,81	-1,423	2,026	0,391	0,625
12.4.2006	15	3 776,94	-0,313	0,098	0,381	0,618
13.4.2006	16	3 779,94	0,079	0,006	0,366	0,605
18.4.2006	17	3 770,79	-0,242	0,059	0,355	0,596
19.4.2006	18	3 820,96	1,330	1,770	0,447	0,669
20.4.2006	19	3 860,00	1,022	1,044	0,492	0,701
21.4.2006	20	3 888,46	0,737	0,544	0,505	0,711
24.4.2006	21	3 862,27	-0,674	0,454	0,512	0,716
25.4.2006	22	3 871,09	0,228	0,052	0,495	0,703

Pokud bychom chtěli předpovědět roční volatilitu míry zisku indexu PX v daném časovém okamžiku přes 252 budoucích obchodních dnů na burze, pak mluvíme o tzv. časově agregované volatilitě a pro její výpočet použijeme aproximaci pomocí předpovědi denní volatility v tomto časovém okamžiku (tj. pro příští obchodní den)

$$\sigma_{roč} = \sqrt{252} \cdot \sigma_{denní}.$$

V tabulce 3.15 jsou uvedeny předpovědi roční volatility míry zisku indexu PX 50 v daném časovém okamžiku (poslední den obchodování v roce) pro rok následující v rozmezí 1995-2006.

Tabulka 3.15 – předpověď roční volatility míry zisku indexu PX 50 za období 1995-2006

<b>časový okamžik předpovědi</b>	15.12. 1995	31.12. 1996	30.12. 1997	30.12. 1998	30.12. 1999	29.12. 2000
<b>předpověď pro rok</b>	1996	1997	1998	1999	2000	2001
<b>předpověď roční volatility</b>	11,33	12,33	17,73	17,18	16,00	34,73
<b>časový okamžik předpovědi</b>	28.12. 2001	30.12. 2002	30.12. 2003	30.12. 2004	30.12. 2005	
<b>předpověď pro rok</b>	2002	2003	2004	2005	2006	
<b>předpověď roční volatility</b>	17,9	18,37	15,43	15,86	11,03	

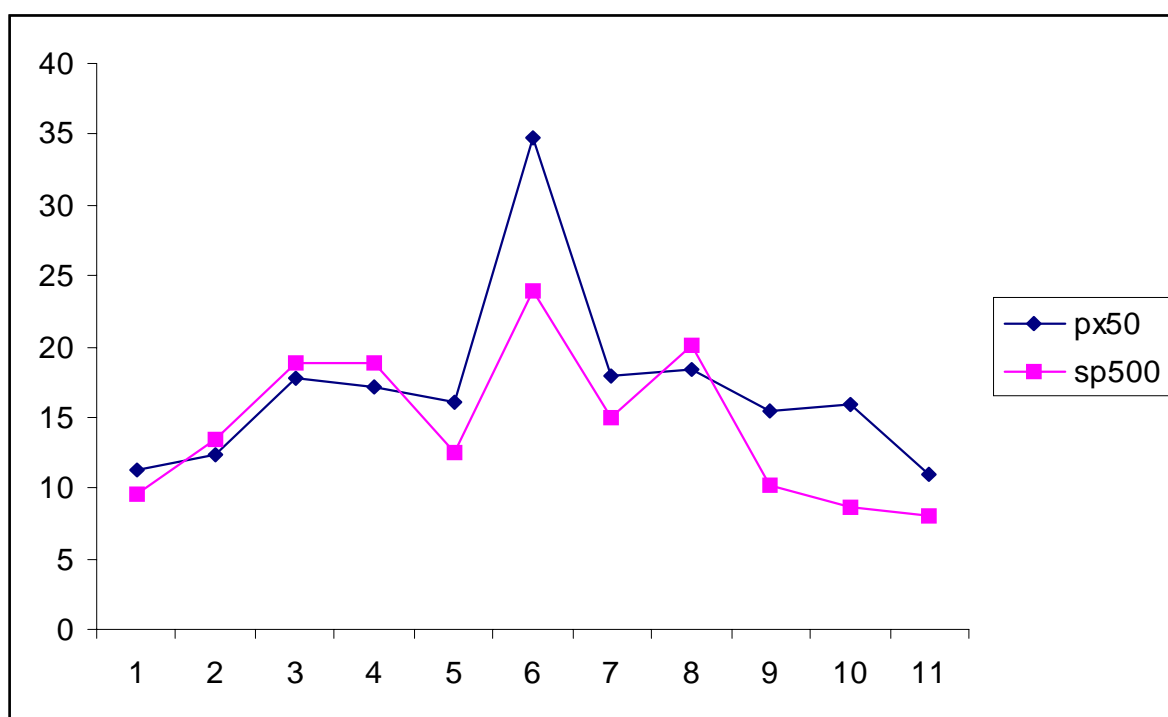
Z tabulky 3.15 vidíme, že předpověď roční volatility míry zisku indexu PX 50 je největší pro rok 2001. Do tohoto roku roste a po tomto roce klesá. Pro letošní rok, tj. rok 2006, je předpověď roční volatility míry zisku indexu PX 50 nejmenší za celé sledované období.

Podívejme se nyní na předpovědi roční volatility míry zisku indexu Standard&Poor's 500. Tyto předpovědi jsou uvedeny v tabulce 3.16. Jde opět o předpovědi v daném časovém okamžiku, a to v poslední den obchodování v roce pro rok následující v rozmezí let 1995-2006.

*Tabulka 3.16 – předpověď roční volatility míry zisku indexu Standard&Poor's 500 za období 1995-2006*

<b>časový okamžik předpovědi</b>	29.12. 1995	31.12. 1996	31.12. 1997	31.12. 1998	31.12. 1999	29.12. 2000
<b>předpověď pro rok</b>	1996	1997	1998	1999	2000	2001
<b>předpověď roční volatility</b>	9,53	13,39	18,77	18,79	12,50	23,88
<b>časový okamžik předpovědi</b>	31.12. 2001	31.12. 2002	31.12. 2003	31.12. 2004	30.12. 2005	
<b>předpověď pro rok</b>	2002	2003	2004	2005	2006	
<b>předpověď roční volatility</b>	14,93	20,07	10,26	8,65	7,96	

Porovnáme-li předpověď roční volatility míry zisku indexu PX 50 (viz tabulka 3.15) a předpověď roční volatility míry zisku indexu Standard&Poor's 500 (viz tabulka 3.16), zjistíme, že se chovají z hlediska monotonií přibližně stejně, tzn. pokud u jednoho indexu předpověď roční volatility míry zisku roste, roste i u druhého a naopak pokud klesá, klesá i u druhého.



Obr.3.15 – Vývoj předpovědi roční volatility míry zisku indexu PX 50 a indexu Standard&Poor's 500 v období 1995 – 2005

Z obr.3.15 vidíme, že předpovědi roční volatility míry zisku indexu Standard&Poor's 500 jsou kolísavější než u míry zisku indexu PX 50. Můžeme tudíž konstatovat, že česká burza je proti americké stabilnější.

## Závěr

Téma disertační práce *Prognostické modely v oblasti modelování finančních časových řad* bylo zvoleno pro svoji aktuálnost. V současné době časové řady ukazatelů finančního trhu začínají mít jistou informační schopnost. Jedním z těchto ukazatelů je i index PX 50. Jedná se o ukazatel vývoje kurzů akcií akciového trhu v České republice. Cílem předložené práce bylo určit model, který by co nejlépe popisoval chování indexu PX 50 a na jehož základě by se mohly dělat prognózy dalšího vývoje tohoto indexu. Za tímto účelem byly řešeny dílčí úlohy, a to určení vhodného modelu popisující závislost indexu PX 50 a oborových indexů, určení vhodného modelu popisující závislost indexu PX 50 a indexu Dow Jones Euro Stoxx 50 (index eurozóny), nalezení nejvhodnějšího modelu pro míru zisku indexu PX 50, určení jednodenních předpovědí volatility míry zisku indexu PX 50 a indexu Standard&Poor's 500 po teroristickém útoku 11.9.2001, určení jednodenních předpovědí volatility míry zisku indexu PX 50, indexu Standard&Poor's 500 a indexu Dow Jones Euro Stoxx 50 pro současné období (je duben 2006) a určení předpovědí roční volatility míry zisku indexu PX 50 a indexu Standard&Poor's 500 za období 1995 – 2006. Při zpracování byly vzaty na zřetel faktory, které by mohly ovlivnit vývoj tohoto indexu.

Při zkoumání závislosti indexu PX 50 na jiných finančních časových řadách byla využita teorie modelů s přenosovou funkcí. Tyto modely umožňují modelovat celou řadu zajímavých vztahů mezi časovými řadami, které bychom jinak velmi těžko modelovali jiným způsobem. Modely s přenosovou funkcí je vhodné používat v takových situacích, kdy jednoznačně chceme zdůraznit a modelem postihnout lineární závislost mezi časovými řadami.

Jako první byl zkoumán vztah indexu PX 50 a oborových indexů, jednak během období 1996 až 1999 a jednak v období 2000 až 2006 (konec února, kdy oborové indexy byly zrušeny). Na základě statistické analýzy se zjistilo, že závislost mezi výstupem (index PX 50) a vstupními řadami (oborové indexy) je pouze ve stejném časovém okamžiku a tudíž pro analýzu stačila lineární regrese. Byl vytvořen regresní model s dosti vysokým koeficientem determinace a nízkou směrodatnou odchylkou reziduí. Za pomoci statistických metod se potvrdilo, že na vysvětlení indexu PX 50 se v letech 1996 až 1999 nejvíce podílely oborové indexy BI13 (doprava a spoje), BI12 (energetika), BI15



(peněžnictví) a BI07 (chemický, farmaceutický a gumárenský průmysl). Zajímavé na tomto výsledku je, že indexy BI13, BI12, BI15 a BI07 patřily mezi indexy, které byly zrušeny mezi posledními (v rozmezí února 2005 až února 2006). Další zajímavostí je, že společnosti podílející se na tvorbě indexu PX k únoru 2006 reprezentují převážně obory jako je energetika, chemický, farmaceutický a gumárenský průmysl a peněžnictví.

V případě období 2000 až 2006 model lineární regrese, kdy se řada indexu PX 50 vysvětlovala pomocí zbývajících oborových indexů ( BI12 – energetika, BI07 - chemický, farmaceutický a gumárenský průmysl, BI16 – služby), vychází špatně, koeficient determinace je nízký. Nakonec došlo k 28. únoru 2006 ke zrušení těchto zbývajících oborových indexů.

Další položená otázka byla, zda vývoj indexu PX 50 je ovlivňován vývojem indexu eurozóny, tj. indexem Dow Jones EURO STOXX 50. Pro analýzu byl zvolen opět model s přenosovou funkcí. Data se uvažovala od 2.1.2003 do 22.11.2005, tedy za poslední tři roky. Opět se zjistilo, že vývoj indexu PX 50 není ovlivněn vývojem indexu Dow Jones EURO STOXX 50 s časovým zpožděním, nýbrž ve stejném časovém okamžiku. Přesto byl vytvořen model s přenosovou funkcí, kdy vstupní řada byla řada indexu Dow Jones EURO STOXX 50 a výstupní řada byla řada PX 50. Tento model nebyl vhodný, což bylo způsobeno řadou faktorů. Např. modely ARIMA pro výstupní řadu PX 50 a vstupní řadu Euroindexu nebyly uspokojivé, v různých úsecích jednotlivých diferencovaných řad byla různá variabilita, v modelu se vyskytovala příliš velká zpoždění atd. Z těchto důvodů nemělo smysl na základě tohoto modelu provádět předpovědi pro burzovní index PX 50 na základě vývoje indexu Dow Jones EURO STOXX 50. Dále byl vytvořen klasický regresní model mezi těmito dvěma řadami, lomený regresní model a posléze model ze dvou nezávislých regresí. V posledním případě vycházel koeficient determinace nejlépe. Při modelování se ukázalo, že důležitou roli hrají pravidelné burzovní cykly, které do značné míry ovlivňují tvorbu modelu.

Vývoj indexu PX 50 je ovlivňován řadou událostí. Jednou z nich byl i teroristický útok na World Trade Center v New Yorku 11.9.2001. V práci byl zkoumán dopad této události na vývoj volatility jak míry zisku indexu PX 50, tak indexu Standard & Poor's. Potvrdilo se, že tato událost měla stejný dopad na vývoj volatility míry zisku obou indexů.

V rámci analýzy byl určen i obecný model volatility pro míru zisku indexu PX 50 a předpovědi volatility. Vzhledem k tomu, že předpověď volatility míry zisku je důležitou součástí některých metod souvisejících s regulací finančních rizik, byly vytvořeny jednak jednodenní předpovědi volatility míry zisku indexu PX 50, Standard&Poor's 500 a Dow Jones EURO STOXX 50 pro současnost (je duben 2006), a jednak předpovědi volatility roční míry zisku indexu PX 50 a indexu Standard&Poor's 500 za období 1995 – 2006. Dospělo se k závěru, že během měsíců březen a duben roku 2006 jednodenní předpovědi volatility míry zisku indexu PX 50, Standard&Poor's 500 a Dow Jones EURO STOXX 50 se chovaly stejně, tzn. rostly. Tento výsledek koresponduje se skutečností, že závislost mezi indexem PX 50 a indexem Dow Jones EURO STOXX 50 je ve stejném časovém okamžiku. Co se týká předpovědi roční volatility míry zisku indexu PX 50 a indexu Standard&Poor's 500 během období 1995-2006, je největší v obou případech pro rok 2001. Do tohoto roku roste a po tomto roce klesá. Pro letošní rok, tj. rok 2006, je předpověď roční volatility míry zisku indexu PX 50 a indexu Standard&Poor's 500 nejmenší za celé sledované období. Z hlediska růstu či poklesu se předpovědi roční volatility míry zisku u obou indexů chovají stejně, přičemž růst či pokles předpovědi roční volatility míry zisku indexu Standard&Poor's 500 je kolísavější.

Dosažené výsledky reflektují skutečný vývoj indexu PX 50, který zachycuje relativní změny v ceně portfolia akcií. Výsledky disertační práce mohou být významným nástrojem pro podporu rozhodování (v kombinaci s ostatními metodami, konkrétně s akciovými analýzami) odhadu dalšího vývoje indexu PX. Vzhledem k tomu, že bez dobře fungujících burz by neexistoval kapitálový trh, který má za úkol shromažďovat volný kapitál a orientovat jej tam, kde ho lze co možná nejúčelněji využít, je důležitost znalosti dalšího vývoje tohoto indexu zřejmá. Na metodické postupy, které byly aplikovány, je možné navázat a po příslušné adaptaci je využít k analýzám jiných světových burzovních indexů, které do značné míry ovlivňují vývoj indexu PX 50.

# **PŘÍLOHA**

## Výstup z programu SCA – Konstrukce modelu PX 50

```

THE PC SCA STATISTICAL SYSTEM ( Release 5.2-Professional )

COPYRIGHT 1985-1998, SCIENTIFIC COMPUTING ASSOCIATES. ALL RIGHTS RESERVED
SCA PRODUCT IDENTIFICATION: UTS & EXPERT (GSA OPTIONAL)
SCA SOFTWARE IDENTIFICATION: UNI.ECONOMICS ( 10528 )
RELEASE DATE: 9/ 1/98

SIZE OF WORKSPACE IS 99000 SINGLE PRECISION WORDS
DATE -- 2/12/** TIME -- 13: 0:38
--

WORKSPACE NOVAR 600.

THE WORKSPACE IS COMPRESSED
THE MAXIMUM NUMBER OF VARIABLES IS EXPANDED TO 608
--

INPUT VARIABLE PX50, Euroindx. FILE 'C:\Petraskova\Indexy.txt'.

PX50 , A 744 BY 1 VARIABLE, IS STORED IN THE WORKSPACE
EUROINDX, A 744 BY 1 VARIABLE, IS STORED IN THE WORKSPACE
--

IDENTIFY VARIABLE PX50.

NAME OF THE SERIES . . . . . PX50
TIME PERIOD ANALYZED . . . . . 1 TO 744
MEAN OF THE (DIFFERENCED) SERIES . . . 861.3291
STANDARD DEVIATION OF THE SERIES . . . 288.7226
T-VALUE OF MEAN (AGAINST ZERO) . . . . 81.3720

AUTOCORRELATIONS

1- 12 1.00 .99 .99 .98 .98 .98 .97 .97 .96 .96 .95 .95
ST.E. .04 .06 .08 .10 .11 .12 .13 .14 .15 .16 .16 .17
Q 741 1477 2207 2933 3653 4368 5078 5782 6481 7174 7861 8543

13- 24 .94 .94 .93 .93 .93 .92 .92 .91 .91 .91 .90 .90
ST.E. .18 .19 .19 .20 .20 .21 .21 .22 .22 .23 .23 .24
Q 9219 988910554112141186912519131641380414440150711569816320

-1.0 -.8 -.6 -.4 -.2 .0 .2 .4 .6 .8 1.0
+-----+-----+-----+-----+-----+-----+-----+-----+
I
1 1.00 + IX+XXXXXXXXXXXXXXXXXXXXXXXXXXXX
2 .99 + IXX+XXXXXXXXXXXXXXXXXXXXXXXXXXXX
3 .99 + IXXX+XXXXXXXXXXXXXXXXXXXXXXXXXXXX
4 .98 + IXXXX+XXXXXXXXXXXXXXXXXXXXXXXXXXXX
5 .98 + IXXXXX+XXXXXXXXXXXXXXXXXXXXXXXXXXXX
6 .98 + IXXXXXX+XXXXXXXXXXXXXXXXXXXXXXXXXXXX
7 .97 + IXXXXXX+XXXXXXXXXXXXXXXXXXXXXXXXXXXX
8 .97 + IXXXXXXX+XXXXXXXXXXXXXXXXXXXXXXXXXXXX
9 .96 + IXXXXXXX+XXXXXXXXXXXXXXXXXXXXXXXXXXXX
10 .96 + IXXXXXXX+XXXXXXXXXXXXXXXXXXXXXXXXXXXX
11 .95 + IXXXXXXX+XXXXXXXXXXXXXXXXXXXXXXXXXXXX
12 .95 + IXXXXXXX+XXXXXXXXXXXXXXXXXXXXXXXXXXXX
13 .94 + IXXXXXXX+XXXXXXXXXXXXXXXXXXXXXXXXXXXX
14 .94 + IXXXXXXX+XXXXXXXXXXXXXXXXXXXXXXXXXXXX
15 .93 + IXXXXXXX+XXXXXXXXXXXXXXXXXXXXXXXXXXXX
16 .93 + IXXXXXXX+XXXXXXXXXXXXXXXXXXXXXXXXXXXX
17 .93 + IXXXXXXX+XXXXXXXXXXXXXXXXXXXXXXXXXXXX
18 .92 + IXXXXXXX+XXXXXXXXXXXXXXXXXXXXXXXXXXXX
19 .92 + IXXXXXXX+XXXXXXXXXXXXXXXXXXXXXXXXXXXX
20 .91 + IXXXXXXX+XXXXXXXXXXXXXXXXXXXXXXXXXXXX

```

```

21 .91 + IXXXXXXXXXX+XXXXXXXXXXXX
22 .91 + IXXXXXXXXXX+XXXXXXXXXXXX
23 .90 + IXXXXXXXXXX+XXXXXXXXXXXX
24 .90 + IXXXXXXXXXX+XXXXXXXXXXXX

```

PARTIAL AUTOCORRELATIONS

```

1- 12 1.00 -.04 .01 .01 -.00 -.00 -.03 -.00 -.01 -.02 -.01 -.02
ST.E. .04 .04 .04 .04 .04 .04 .04 .04 .04 .04 .04 .04 .04

13- 24 .02 .02 .00 .00 .02 .01 .00 .00 .01 .01 .01 -.00
ST.E. .04 .04 .04 .04 .04 .04 .04 .04 .04 .04 .04 .04 .04

```

```

-1.0 -.8 -.6 -.4 -.2 .0 .2 .4 .6 .8 1.0
+-----+-----+-----+-----+-----+

```

```

1 1.00 + IXXXXXXXXXXXXXXXXXXXXXXXXXXXX
2 -.04 +XI +
3 .01 + I +
4 .01 + I +
5 .00 + I +
6 .00 + I +
7 -.03 +XI +
8 .00 + I +
9 -.01 + I +
10 -.02 +XI +
11 -.01 + I +
12 -.02 + I +
13 .02 + IX+
14 .02 + I +
15 .00 + I +
16 .00 + I +
17 .02 + IX+
18 .01 + I +
19 .00 + I +
20 .00 + I +
21 .01 + I +
22 .01 + I +
23 .01 + I +
24 .00 + I +
--

```

IDENTIFY VARIABLE PX50. DFORDER 1

```

DIFFERENCE ORDERS . . . . . (1-B )
NAME OF THE SERIES . . . . . PX50
TIME PERIOD ANALYZED . . . . . 1 TO 744
MEAN OF THE (DIFFERENCED) SERIES . . . 1.2295
STANDARD DEVIATION OF THE SERIES . . . 9.6360
T-VALUE OF MEAN (AGAINST ZERO) . . . . 3.4779

```

AUTOCORRELATIONS

```

1- 12 .13 -.04 -.05 -.01 .10 .08 -.10 -.05 .14 -.03 -.02 -.02
ST.E. .04 .04 .04 .04 .04 .04 .04 .04 .04 .04 .04 .04 .04
Q 13.5 14.8 16.4 16.4 23.4 28.6 35.8 37.4 51.3 51.8 52.3 52.7

13- 24 -.09 -.01 -.04 -.08 -.01 .06 -.01 -.07 .04 .07 -.01 -.05
ST.E. .04 .04 .04 .04 .04 .04 .04 .04 .04 .04 .04 .04 .04
Q 59.4 59.5 60.7 66.1 66.1 68.8 68.8 72.1 73.2 76.6 76.7 78.4

```

```

      -1.0  -.8  -.6  -.4  -.2  .0  .2  .4  .6  .8  1.0
      +-----+-----+-----+-----+-----+
                I
1      .13          + IX+X
2     -.04         +XI +
3     -.05         +XI +
4     -.01         + I +
5      .10         + IXX
6      .08         + IXX
7     -.10         XXI +
8     -.05         +XI +
9      .14         + IX+X
10    -.03         +XI +
11    -.02         +XI +
12    -.02         +XI +
13    -.09         XXI +
14    -.01         + I +
15    -.04         +XI +
16    -.08         XXI +
17    -.01         + I +
18     .06         + IX+
19    -.01         + I +
20    -.07         XXI +
21     .04         + IX+
22     .07         + IXX
23    -.01         + I +
24    -.05         +XI +

```

PARTIAL AUTOCORRELATIONS

```

1- 12      .13  -.06  -.03  .00  .10  .06  -.11  -.01  .15  -.09  -.02  .00
ST.E.      .04  .04  .04  .04  .04  .04  .04  .04  .04  .04  .04  .04

13- 24     -.08  -.02  -.07  -.04  .01  .04  -.01  -.07  .08  .08  -.08  -.03
ST.E.      .04  .04  .04  .04  .04  .04  .04  .04  .04  .04  .04  .04

```

```

      -1.0  -.8  -.6  -.4  -.2  .0  .2  .4  .6  .8  1.0
      +-----+-----+-----+-----+-----+
                I
1      .13          + IX+X
2     -.06         XXI +
3     -.03         +XI +
4      .00         + I +
5      .10         + IXX
6      .06         + IX+
7     -.11         X+XI +
8     -.01         + I +
9      .15         + IX+XX
10    -.09         XXI +
11    -.02         + I +
12     .00         + I +
13    -.08         XXI +
14    -.02         + I +
15    -.07         XXI +
16    -.04         +XI +
17     .01         + I +
18     .04         + IX+
19    -.01         + I +
20    -.07         XXI +
21     .08         + IXX
22     .08         + IXX
23    -.08         XXI +
24    -.03         +XI +

```

IDENTIFY VARIABLE Euroindx.

```

NAME OF THE SERIES . . . . . EUROINDX
TIME PERIOD ANALYZED . . . . . 1 TO 744

```

MEAN OF THE (DIFFERENCED) SERIES . . . 2786.1050  
 STANDARD DEVIATION OF THE SERIES . . . 336.9312  
 T-VALUE OF MEAN (AGAINST ZERO) . . . 225.5500

AUTOCORRELATIONS

1- 12	.99	.99	.98	.97	.97	.96	.96	.95	.95	.94	.94	.93
ST.E.	.04	.06	.08	.10	.11	.12	.13	.14	.15	.16	.16	.17
Q	737	1466	2185	2897	3602	4299	4988	5672	6349	7020	7686	8344
13- 24	.93	.92	.91	.91	.90	.90	.89	.89	.88	.88	.87	.87
ST.E.	.18	.18	.19	.19	.20	.21	.21	.22	.22	.23	.23	.23
Q	8996	9640	10277	10907	11530	12148	12759	13364	13962	14554	15140	15719

	-1.0	-.8	-.6	-.4	-.2	.0	.2	.4	.6	.8	1.0	
	+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+											
	I											
1 .99												
2 .99												
3 .98												
4 .97												
5 .97												
6 .96												
7 .96												
8 .95												
9 .95												
10 .94												
11 .94												
12 .93												
13 .93												
14 .92												
15 .91												
16 .91												
17 .90												
18 .90												
19 .89												
20 .89												
21 .88												
22 .88												
23 .87												
24 .87												

PARTIAL AUTOCORRELATIONS

1- 12	.99	.00	.00	.04	-.01	.02	.01	.06	-.02	.00	.02	-.04
ST.E.	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04
13- 24	-.01	-.04	.00	.01	.01	.03	.01	-.05	.01	-.02	.02	-.02
ST.E.	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04

	-1.0	-.8	-.6	-.4	-.2	.0	.2	.4	.6	.8	1.0	
	+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+											
	I											
1 .99												
2 .00												
3 .00												
4 .04												
5 -.01												
6 .02												
7 .01												
8 .06												
9 -.02												
10 .00												
11 .02												
12 -.04												
13 -.01												
14 -.04												
15 .00												
16 .01												

```

17 .01 + I +
18 .03 + IX+
19 .01 + I +
20 -.05 +XI +
21 .01 + I +
22 -.02 +XI +
23 .02 + IX+
24 -.02 + I +
--

```

IDENTIFY VARIABLE Euroindx. DFORDER 1

```

1
DIFFERENCE ORDERS. . . . . (1-B )
NAME OF THE SERIES . . . . . EUROINDX
TIME PERIOD ANALYZED . . . . . 1 TO 744
MEAN OF THE (DIFFERENCED) SERIES . . . 1.2494
STANDARD DEVIATION OF THE SERIES . . . 28.5731
T-VALUE OF MEAN (AGAINST ZERO) . . . . 1.1919

```

AUTOCORRELATIONS

```

1- 12   -.02  .03  -.07  .01  -.04  -.03  -.11  .09  -.02  -.04  .03  .01
ST.E.   .04  .04  .04  .04  .04  .04  .04  .04  .04  .04  .04  .04
Q       .2   .8  4.5  4.6  5.6  6.3 15.5 21.3 21.5 22.5 23.2 23.3

13- 24   .05  .01  -.03  -.06  -.01  -.03  .03  .01  .02  -.00  .01  -.07
ST.E.   .04  .04  .04  .04  .04  .04  .04  .04  .04  .04  .04  .04
Q      25.2 25.2 26.0 28.4 28.5 29.3 30.0 30.1 30.3 30.3 30.4 33.6

```

```

-1.0  -.8  -.6  -.4  -.2  .0  .2  .4  .6  .8  1.0
+-----+-----+-----+-----+-----+-----+-----+-----+-----+

```

```

I
1  -.02 + I +
2  .03 + IX+
3  -.07 XXI +
4  .01 + I +
5  -.04 +XI +
6  -.03 +XI +
7  -.11 X+XI +
8  .09 + IXX
9  -.02 + I +
10 -.04 +XI +
11 .03 + IX+
12 .01 + I +
13 .05 + IX+
14 .01 + I +
15 -.03 +XI +
16 -.06 +XI +
17 -.01 + I +
18 -.03 +XI +
19 .03 + IX+
20 .01 + I +
21 .02 + I +
22 .00 + I +
23 .01 + I +
24 -.07 XXI +

```

PARTIAL AUTOCORRELATIONS

```

1- 12   -.02  .03  -.07  .01  -.03  -.04  -.11  .08  -.01  -.06  .04  .00
ST.E.   .04  .04  .04  .04  .04  .04  .04  .04  .04  .04  .04  .04

13- 24   .04  .01  -.02  -.06  -.01  -.02  .02  .02  .01  -.01  .00  -.06
ST.E.   .04  .04  .04  .04  .04  .04  .04  .04  .04  .04  .04  .04

```

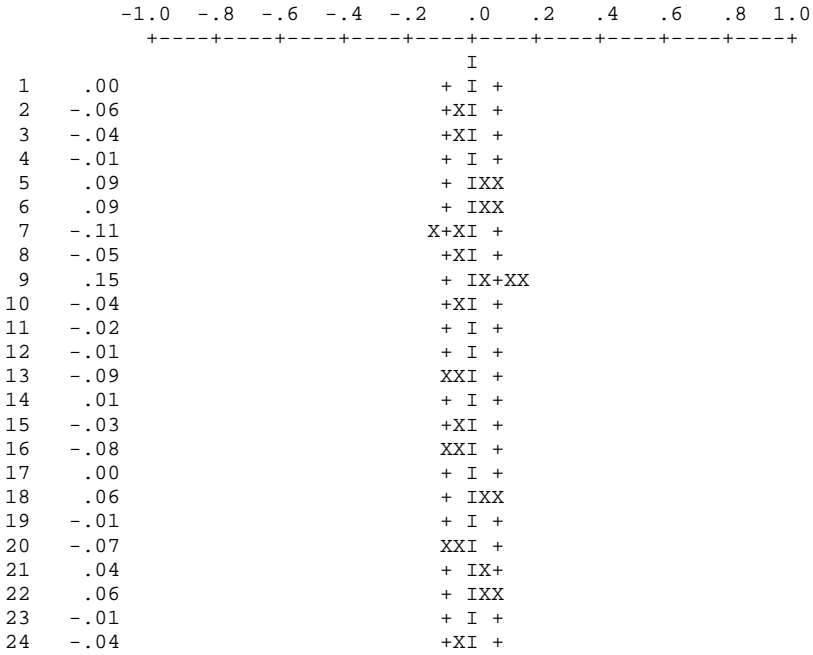




STANDARD DEVIATION OF THE SERIES . . . 9.5551  
T-VALUE OF MEAN (AGAINST ZERO) . . . 2.9648

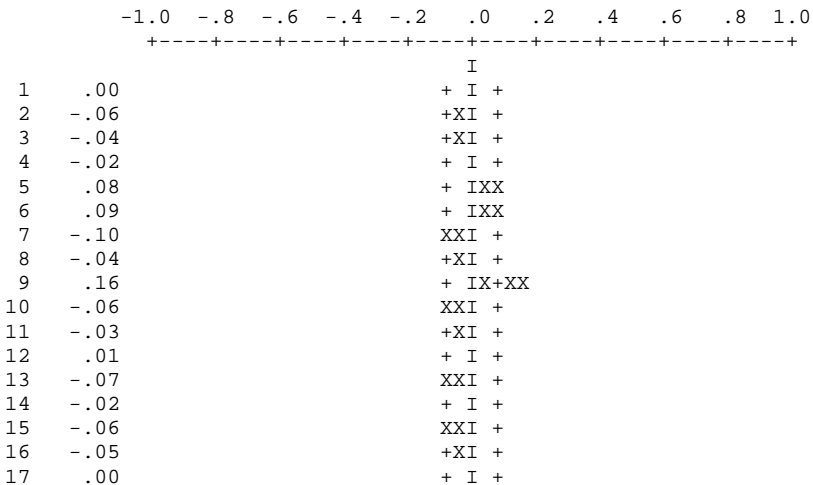
AUTOCORRELATIONS

1- 12	-.00	-.06	-.04	-.01	.09	.09	-.11	-.05	.15	-.04	-.02	-.01
ST.E.	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04
Q	.0	2.4	3.6	3.8	9.6	15.3	23.8	26.0	43.6	45.1	45.3	45.3
13- 24	-.09	.01	-.03	-.08	-.00	.06	-.01	-.07	.04	.06	-.01	-.04
ST.E.	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04
Q	51.9	52.1	52.7	57.7	57.7	60.7	60.8	64.8	66.0	69.1	69.3	70.5



PARTIAL AUTOCORRELATIONS

1- 12	-.00	-.06	-.04	-.02	.08	.09	-.10	-.04	.16	-.06	-.03	.01
ST.E.	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04
13- 24	-.07	-.02	-.06	-.05	-.00	.04	.01	-.08	.06	.10	-.06	-.04
ST.E.	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04



```

18 .04 + IX+
19 .01 + I +
20 -.08 XXI +
21 .06 + IX+
22 .10 + IXX
23 -.06 +XI +
24 -.04 +XI +

```

```

TSMODEL NAME Index. MODEL @
(1- phil *B)Euroindx(1) = NOISE.

```

```
--
```

```

ESTIM MODEL Index. HOLD @
RESIDUALS(resY1),FITTED(fity1)

```

```
THE FOLLOWING ANALYSIS IS BASED ON TIME SPAN 1 THRU 744
```

```

NONLINEAR ESTIMATION TERMINATED DUE TO:
RELATIVE CHANGE IN (OBJECTIVE FUNCTION)**0.5 LESS THAN .1000D-02

```

```
SUMMARY FOR UNIVARIATE TIME SERIES MODEL -- INDEX
```

```

-----
VARIABLE TYPE OF ORIGINAL DIFFERENCING
VARIABLE OR CENTERED
1
EUROINDX RANDOM ORIGINAL (1-B )
-----

```

PARAMETER LABEL	VARIABLE NAME	NUM./ DENOM.	FACTOR	ORDER	CONS- TRAIT	VALUE	STD ERROR	T VALUE	
1	PHI1	EUROINDX	AR	1	1	NONE	-.0156	.0367	-.43

```

EFFECTIVE NUMBER OF OBSERVATIONS . . . . . 742
R-SQUARE . . . . . .993
RESIDUAL STANDARD ERROR. . . . . .286067E+02
--

```

```
IDENTIFY VARIABLE RESY1.
```

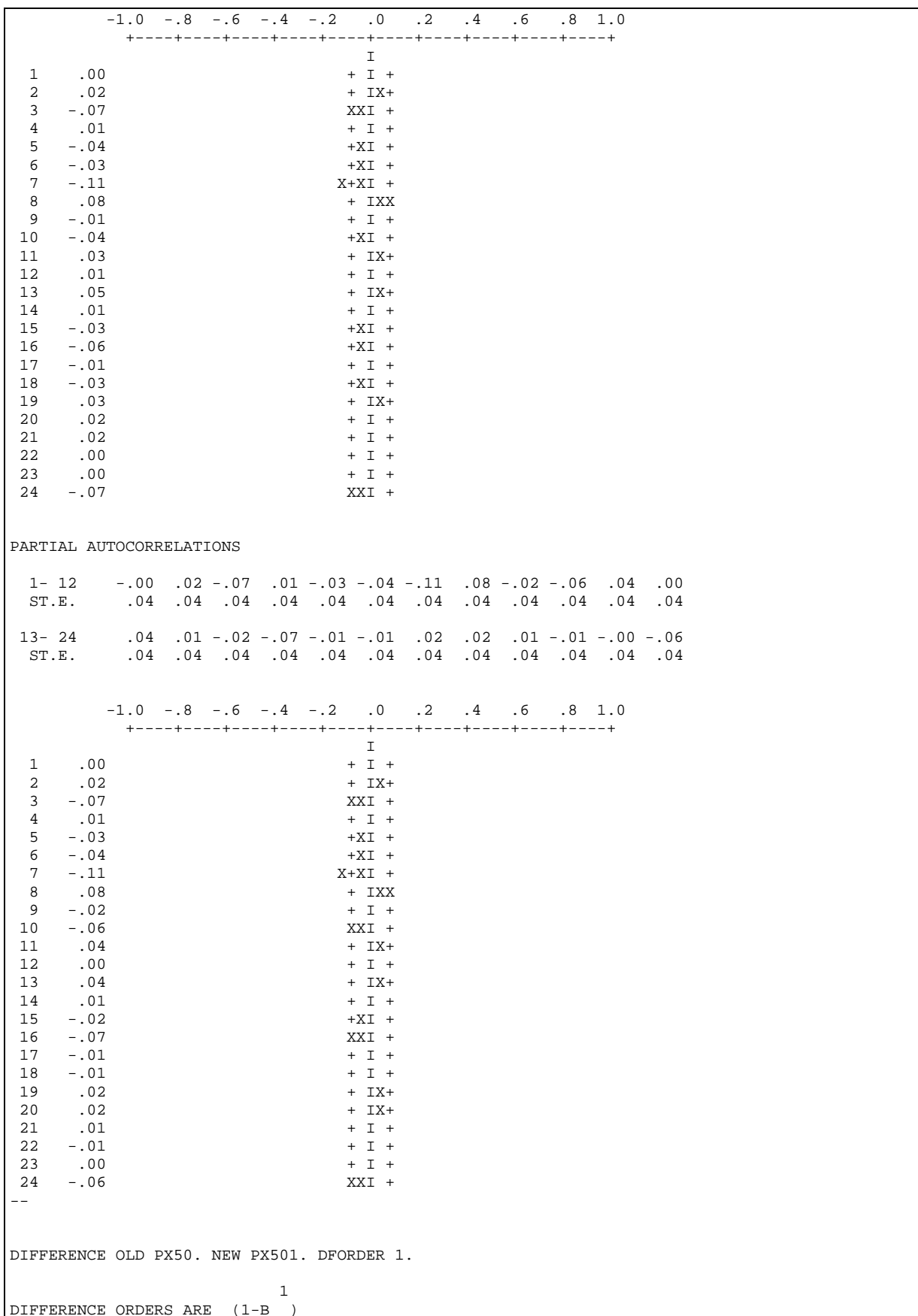
```

NAME OF THE SERIES . . . . . RESY1
TIME PERIOD ANALYZED . . . . . 3 TO 744
MEAN OF THE (DIFFERENCED) SERIES . . . . . 1.2976
STANDARD DEVIATION OF THE SERIES . . . . . 28.5773
T-VALUE OF MEAN (AGAINST ZERO) . . . . . 1.2368

```

```
AUTOCORRELATIONS
```

1- 12	-.00	.02	-.07	.01	-.04	-.03	-.11	.08	-.01	-.04	.03	.01
ST.E.	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04
Q	.0	.4	4.4	4.4	5.5	6.2	15.3	20.7	20.9	22.0	22.7	22.7
13- 24	.05	.01	-.03	-.06	-.01	-.03	.03	.02	.02	-.00	.00	-.07
ST.E.	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04
Q	24.5	24.6	25.4	28.1	28.2	28.9	29.7	29.9	30.1	30.1	30.1	33.6



SERIES PX50 IS DIFFERENCED, THE RESULT IS STORED IN VARIABLE PX501  
 SERIES PX501 HAS 744 ENTRIES

DIFFERENCE OLD Euroindx. NEW Euroind1. DFORDER 1.

1  
 DIFFERENCE ORDERS ARE (1-B )  
 SERIES EUROINDX IS DIFFERENCED, THE RESULT IS STORED IN VARIABLE EUROIND1  
 SERIES EUROIND1 HAS 744 ENTRIES

CCF VARIABLE PX501, Euroind1.

TIME PERIOD ANALYZED . . . . .	2	TO	744	
NAMES OF THE SERIES . . . . .		PX501		EUROIND1
EFFECTIVE NUMBER OF OBSERVATIONS . . .		743		743
STANDARD DEVIATION OF THE SERIES . . .		9.6360		28.5731
MEAN OF THE (DIFFERENCED) SERIES . . .		1.2295		1.2494
STANDARD DEVIATION OF THE MEAN . . . .		.3535		1.0482
T-VALUE OF MEAN (AGAINST ZERO) . . . .		3.4779		1.1919

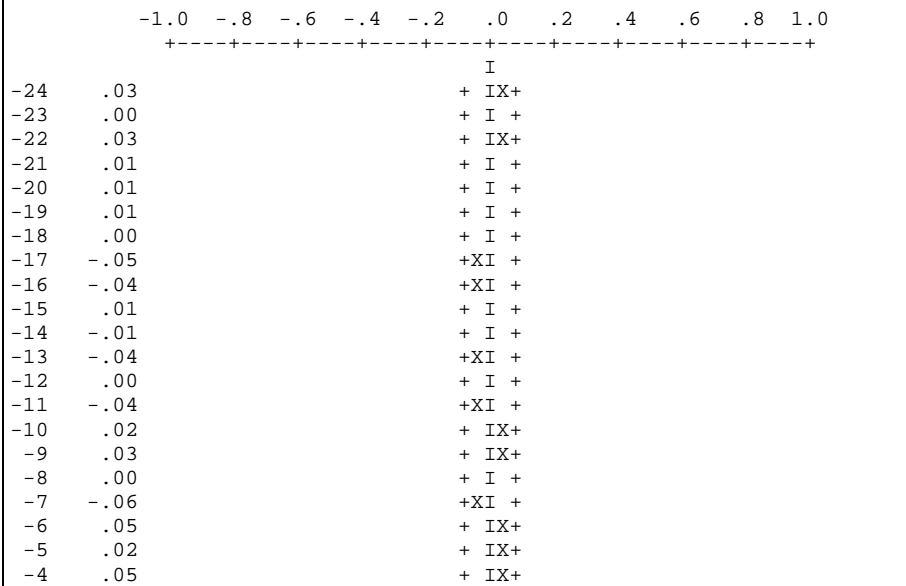
CORRELATION BETWEEN EUROIND1 AND PX501 IS .26

CROSS CORRELATION BETWEEN PX501(T) AND EUROIND1(T-L)

1- 12	.06	.02	-.00	.03	.05	.00	-.02	-.04	-.04	-.00	.03	.02
ST.E.	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04
13- 24	-.06	-.01	-.02	-.01	.04	.04	.02	.04	-.01	.03	-.06	-.03
ST.E.	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04

CROSS CORRELATION BETWEEN EUROIND1(T) AND PX501(T-L)

1- 12	.03	.00	-.03	.05	.02	.05	-.06	.00	.03	.02	-.04	-.00
ST.E.	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04
13- 24	-.04	-.01	.01	-.04	-.05	-.00	.01	.01	.01	.03	.00	.03
ST.E.	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04



```

-3  -.03      +XI +
-2  .00       + I +
-1  .03       + IX+
 0  .26       + IX+XXXX
 1  .06       + IXX
 2  .02       + I +
 3  .00       + I +
 4  .03       + IX+
 5  .05       + IX+
 6  .00       + I +
 7  -.02      + I +
 8  -.04      +XI +
 9  -.04      +XI +
10  .00       + I +
11  .03       + IX+
12  .02       + I +
13  -.06      +XI +
14  -.01      + I +
15  -.02      + I +
16  -.01      + I +
17  .04       + IX+
18  .04       + IX+
19  .02       + I +
20  .04       + IX+
21  -.01      + I +
22  .03       + IX+
23  -.06      +XI +
24  -.03      +XI +
--

```

```

TSMODEL NAME PXEur. MODEL @
PX50(1)=(0 to 10;V0 to V10)Euroindx(1)+NOISE
--

```

```

ESTIM MODEL PXEur. HOLD @
RESIDUALS(resY),FITTED(fitY),DISTURBANCE(NT).

```

THE FOLLOWING ANALYSIS IS BASED ON TIME SPAN 1 THRU 744

NONLINEAR ESTIMATION TERMINATED DUE TO:  
RELATIVE CHANGE IN (OBJECTIVE FUNCTION)\*\*0.5 LESS THAN .1000D-02

SUMMARY FOR UNIVARIATE TIME SERIES MODEL -- PXEUR

```

-----
VARIABLE   TYPE OF   ORIGINAL   DIFFERENCING
           VARIABLE OR CENTERED
           1
PX50       RANDOM   ORIGINAL   (1-B )
           1
EUROINDX   RANDOM   ORIGINAL   (1-B )
-----

```

PARAMETER LABEL	VARIABLE NAME	NUM. / DENOM.	FACTOR	ORDER	CONS- TRAIT	VALUE	STD ERROR	T VALUE	
1	V0	EUROINDX	NUM.	1	0	NONE	.0926	.0123	7.55
2	V1	EUROINDX	NUM.	1	1	NONE	.0245	.0122	2.00
3	V2	EUROINDX	NUM.	1	2	NONE	.0036	.0123	.30
4	V3	EUROINDX	NUM.	1	3	NONE	.0071	.0122	.58
5	V4	EUROINDX	NUM.	1	4	NONE	.0147	.0122	1.21
6	V5	EUROINDX	NUM.	1	5	NONE	.0209	.0122	1.72
7	V6	EUROINDX	NUM.	1	6	NONE	.0064	.0122	.52
8	V7	EUROINDX	NUM.	1	7	NONE	.0068	.0122	.56
9	V8	EUROINDX	NUM.	1	8	NONE	-.0149	.0122	-1.22
10	V9	EUROINDX	NUM.	1	9	NONE	-.0102	.0122	-.84
11	V10	EUROINDX	NUM.	1	10	NONE	.0067	.0122	.55



```

5      .08      + IXX
6      .04      + IX+
7     -.11      X+XI +
8      .00      + I +
9      .16      + IX+XX
10     -.09      XXI +
11     -.02      + I +
12     .00      + I +
13     -.05      +XI +
14     -.02      + I +
15     -.07      XXI +
16     -.04      +XI +
17     .00      + I +
18     .04      + IX+
19     -.02      + I +
20     -.08      XXI +
21     .09      + IXX
22     .06      + IX+
23     -.05      +XI +
24     -.04      +XI +
--

```

IDENTIFY VARIABLE RESY.

```

NAME OF THE SERIES . . . . . RESY
TIME PERIOD ANALYZED . . . . . 12 TO 744
MEAN OF THE (DIFFERENCED) SERIES . . . 1.0124
STANDARD DEVIATION OF THE SERIES . . . 9.3005
T-VALUE OF MEAN (AGAINST ZERO) . . . . 2.9470

```

AUTOCORRELATIONS

```

1- 12      .11 -.05 -.04 -.03 .07 .06 -.10 -.03 .15 -.04 -.03 -.02
ST.E.      .04 .04 .04 .04 .04 .04 .04 .04 .04 .04 .04 .04
Q          9.4 11.4 12.8 13.6 17.5 20.5 27.5 28.1 44.7 46.1 46.7 47.1

13- 24     -.07 .00 -.05 -.09 -.01 .07 -.02 -.08 .04 .06 .01 -.05
ST.E.      .04 .04 .04 .04 .04 .04 .04 .04 .04 .04 .04 .04
Q          50.7 50.7 52.6 58.8 58.8 62.1 62.3 67.1 68.7 71.3 71.3 73.4

```

```

-1.0  -.8  -.6  -.4  -.2  .0  .2  .4  .6  .8  1.0
+-----+-----+-----+-----+-----+-----+
I
1      .11      + IX+X
2     -.05      +XI +
3     -.04      +XI +
4     -.03      +XI +
5      .07      + IXX
6      .06      + IXX
7     -.10      XXI +
8     -.03      +XI +
9      .15      + IX+XX
10     -.04      +XI +
11     -.03      +XI +
12     -.02      +XI +
13     -.07      XXI +
14     .00      + I +
15     -.05      +XI +
16     -.09      XXI +
17     -.01      + I +
18     .07      + IXX
19     -.02      + I +
20     -.08      XXI +
21     .04      + IX+
22     .06      + IX+
23     .01      + I +
24     -.05      +XI +

```



PARTIAL AUTOCORRELATIONS

1- 12	.11	-.07	-.03	-.03	.08	.04	-.11	.00	.16	-.09	-.02	.00
ST.E.	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04
13- 24	-.05	-.02	-.07	-.04	-.00	.04	-.02	-.08	.09	.06	-.05	-.04
ST.E.	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04

-1.0 - .8 - .6 - .4 - .2 .0 .2 .4 .6 .8 1.0  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

		I
1	.11	+ IX+X
2	-.07	XXI +
3	-.03	+XI +
4	-.03	+XI +
5	.08	+ IXX
6	.04	+ IX+
7	-.11	X+XI +
8	.00	+ I +
9	.16	+ IX+XX
10	-.09	XXI +
11	-.02	+ I +
12	.00	+ I +
13	-.05	+XI +
14	-.02	+ I +
15	-.07	XXI +
16	-.04	+XI +
17	.00	+ I +
18	.04	+ IX+
19	-.02	+ I +
20	-.08	XXI +
21	.09	+ IXX
22	.06	+ IX+
23	-.05	+XI +
24	-.04	+XI +

--  
 TSMODEL NAME PXEur1. MODEL @  
 (1-phi1\*B-phi9\*B\*\*9-phi20\*B\*\*20)PX50(1)= (v0)Euroindx(1)+NOISE

--  
 ESTIM MODEL PXEur1. HOLD @  
 RESIDUALS(resY),FITTED(fitY),DISTURBANCE(NT).

THE FOLLOWING ANALYSIS IS BASED ON TIME SPAN 1 THRU 744  
 NONLINEAR ESTIMATION TERMINATED DUE TO:  
 RELATIVE CHANGE IN (OBJECTIVE FUNCTION)\*\*0.5 LESS THAN .1000D-02

SUMMARY FOR UNIVARIATE TIME SERIES MODEL -- PXEUR1

VARIABLE	TYPE OF VARIABLE	ORIGINAL OR CENTERED	DIFFERENCING
PX50	RANDOM	ORIGINAL	1 (1-B )
EUROINDX	RANDOM	ORIGINAL	1 (1-B )

PARAMETER LABEL	VARIABLE NAME	NUM. / DENOM.	FACTOR	ORDER	CONS- TRAIT	VALUE	STD ERROR	T VALUE	
1	V0	EUROINDX	NUM.	1	0	NONE	.0869	.0122	7.13
2	PHI1	PX50	AR	1	1	NONE	.1456	.0352	4.14
3	PHI9	PX50	AR	1	9	NONE	.1479	.0354	4.17
4	PHI20	PX50	AR	1	20	NONE	-.0571	.0357	-1.60

EFFECTIVE NUMBER OF OBSERVATIONS . . . . . 723  
R-SQUARE . . . . . .999  
RESIDUAL STANDARD ERROR. . . . . .926179E+01  
--

IDENTIFY VARIABLE NT.

NAME OF THE SERIES . . . . . NT  
TIME PERIOD ANALYZED . . . . . 22 TO 744  
MEAN OF THE (DIFFERENCED) SERIES . . . . .8025  
STANDARD DEVIATION OF THE SERIES . . . . .9.2270  
T-VALUE OF MEAN (AGAINST ZERO) . . . . .2.3386

AUTOCORRELATIONS

1- 12	.00	-.04	-.04	-.05	.08	.09	-.08	-.04	.01	-.08	.00	-.02
ST.E.	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04
Q	.0	1.2	2.5	4.1	8.4	14.3	19.4	20.7	20.8	25.3	25.3	25.6
13- 24	-.08	.02	-.03	-.07	.00	.02	-.01	-.03	.05	.06	-.00	-.06
ST.E.	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04
Q	30.6	30.8	31.7	35.1	35.2	35.3	35.5	36.3	38.5	41.6	41.6	44.0

-1.0 -.8 -.6 -.4 -.2 .0 .2 .4 .6 .8 1.0  
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

		I
1	.00	+ I +
2	-.04	+XI +
3	-.04	+XI +
4	-.05	+XI +
5	.08	+ IXX
6	.09	+ IXX
7	-.08	XXI +
8	-.04	+XI +
9	.01	+ I +
10	-.08	XXI +
11	.00	+ I +
12	-.02	+ I +
13	-.08	XXI +
14	.02	+ I +
15	-.03	+XI +
16	-.07	XXI +
17	.00	+ I +
18	.02	+ I +
19	-.01	+ I +
20	-.03	+XI +
21	.05	+ IX+
22	.06	+ IXX
23	.00	+ I +
24	-.06	+XI +

PARTIAL AUTOCORRELATIONS

1- 12	.00	-.04	-.04	-.05	.07	.09	-.08	-.03	.02	-.09	-.02	-.02
ST.E.	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04
13- 24	-.07	.00	-.04	-.06	-.01	.02	-.01	-.06	.07	.07	-.03	-.06
ST.E.	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04

-1.0    -.8    -.6    -.4    -.2    .0    .2    .4    .6    .8    1.0  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

		I
1	.00	+ I +
2	-.04	+XI +
3	-.04	+XI +
4	-.05	+XI +
5	.07	+ IXX
6	.09	+ IXX
7	-.08	XXI +
8	-.03	+XI +
9	.02	+ IX+
10	-.09	XXI +
11	-.02	+ I +
12	-.02	+XI +
13	-.07	XXI +
14	.00	+ I +
15	-.04	+XI +
16	-.06	XXI +
17	-.01	+ I +
18	.02	+ I +
19	-.01	+ I +
20	-.06	+XI +
21	.07	+ IXX
22	.07	+ IXX
23	-.03	+XI +
24	-.06	+XI +

IDENTIFY VARIABLE RESY.

NAME OF THE SERIES . . . . .	RESY
TIME PERIOD ANALYZED . . . . .	22 TO 744
MEAN OF THE (DIFFERENCED) SERIES . . . . .	.8025
STANDARD DEVIATION OF THE SERIES . . . . .	9.2270
T-VALUE OF MEAN (AGAINST ZERO) . . . . .	2.3386

AUTOCORRELATIONS

1- 12	.00	-.04	-.04	-.05	.08	.09	-.08	-.04	.01	-.08	.00	-.02
ST.E.	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04
Q	.0	1.2	2.5	4.1	8.4	14.3	19.4	20.7	20.8	25.3	25.3	25.6
13- 24	-.08	.02	-.03	-.07	.00	.02	-.01	-.03	.05	.06	-.00	-.06
ST.E.	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04
Q	30.6	30.8	31.7	35.1	35.2	35.3	35.5	36.3	38.5	41.6	41.6	44.0

-1.0    -.8    -.6    -.4    -.2    .0    .2    .4    .6    .8    1.0  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

		I
1	.00	+ I +
2	-.04	+XI +
3	-.04	+XI +
4	-.05	+XI +
5	.08	+ IXX
6	.09	+ IXX
7	-.08	XXI +
8	-.04	+XI +

```

 9   .01           + I +
10  -.08          XXI +
11   .00           + I +
12  -.02           + I +
13  -.08          XXI +
14   .02           + I +
15  -.03          +XI +
16  -.07          XXI +
17   .00           + I +
18   .02           + I +
19  -.01           + I +
20  -.03          +XI +
21   .05           + IX+
22   .06           + IXX
23   .00           + I +
24  -.06          +XI +

```

PARTIAL AUTOCORRELATIONS

```

1- 12   .00  -.04  -.04  -.05  .07  .09  -.08  -.03  .02  -.09  -.02  -.02
ST.E.   .04  .04  .04  .04  .04  .04  .04  .04  .04  .04  .04  .04

13- 24  -.07  .00  -.04  -.06  -.01  .02  -.01  -.06  .07  .07  -.03  -.06
ST.E.   .04  .04  .04  .04  .04  .04  .04  .04  .04  .04  .04  .04

```

```

      -1.0  -.8  -.6  -.4  -.2  .0  .2  .4  .6  .8  1.0
      +-----+-----+-----+-----+-----+
                I
 1   .00           + I +
 2  -.04          +XI +
 3  -.04          +XI +
 4  -.05          +XI +
 5   .07           + IXX
 6   .09           + IXX
 7  -.08          XXI +
 8  -.03          +XI +
 9   .02           + IX+
10  -.09          XXI +
11  -.02           + I +
12  -.02          +XI +
13  -.07          XXI +
14   .00           + I +
15  -.04          +XI +
16  -.06          XXI +
17  -.01           + I +
18   .02           + I +
19  -.01           + I +
20  -.06          +XI +
21   .07           + IXX
22   .07           + IXX
23  -.03          +XI +
24  -.06          +XI +
--

```

a poslední model

```

TSMODEL NAME PXEur1. MODEL @
(1-phi1*B-phi9*B**9)PX50(1)= (v0)Euroindx(1)+NOISE
--

```

```

ESTIM MODEL PXEur1. HOLD @
RESIDUALS(resY),FITTED(fitY),DISTURBANCE(NT).

```

THE FOLLOWING ANALYSIS IS BASED ON TIME SPAN 1 THRU 744

NONLINEAR ESTIMATION TERMINATED DUE TO:  
RELATIVE CHANGE IN (OBJECTIVE FUNCTION)\*\*0.5 LESS THAN .1000D-02

SUMMARY FOR UNIVARIATE TIME SERIES MODEL -- PXEUR1

VARIABLE	TYPE OF VARIABLE	ORIGINAL OR CENTERED	DIFFERENCING
PX50	RANDOM	ORIGINAL	1 (1-B )
EUROINDX	RANDOM	ORIGINAL	1 (1-B )

PARAMETER LABEL	VARIABLE NAME	NUM./ DENOM.	FACTOR	ORDER	CONS- TRRAINT	VALUE	STD ERROR	T VALUE	
1	V0	EUROINDX	NUM.	1	0	NONE	.0855	.0119	7.17
2	PHI1	PX50	AR	1	1	NONE	.1447	.0349	4.14
3	PHI9	PX50	AR	1	9	NONE	.1482	.0352	4.21

EFFECTIVE NUMBER OF OBSERVATIONS . . . . . 734  
R-SQUARE . . . . . .999  
RESIDUAL STANDARD ERROR. . . . . .921617E+01  
--

IDENTIFY VARIABLE NT.

NAME OF THE SERIES . . . . . NT  
TIME PERIOD ANALYZED . . . . . 11 TO 744  
MEAN OF THE (DIFFERENCED) SERIES . . . . .7338  
STANDARD DEVIATION OF THE SERIES . . . . .9.1869  
T-VALUE OF MEAN (AGAINST ZERO) . . . . .2.1639

AUTOCORRELATIONS

1- 12	.00	-.05	-.04	-.04	.08	.09	-.08	-.04	.01	-.07	.00	-.01
ST.E.	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04
Q	.0	1.6	3.0	4.1	8.8	14.9	20.1	21.5	21.6	25.7	25.7	25.8
13- 24	-.08	.01	-.04	-.06	.01	.02	-.02	-.09	.05	.07	.00	-.05
ST.E.	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04
Q	30.4	30.5	32.0	35.1	35.1	35.4	35.6	41.7	43.3	46.8	46.8	49.0

-1.0 -.8 -.6 -.4 -.2 .0 .2 .4 .6 .8 1.0  
+-----+-----+-----+-----+-----+-----+-----+-----+-----+

	I
1	.00 + I +
2	-.05 +XI +
3	-.04 +XI +
4	-.04 +XI +
5	.08 + IXX
6	.09 + IXX
7	-.08 XXI +
8	-.04 +XI +
9	.01 + I +
10	-.07 XXI +
11	.00 + I +
12	-.01 + I +
13	-.08 XXI +
14	.01 + I +
15	-.04 +XI +
16	-.06 XXI +
17	.01 + I +
18	.02 + IX+
19	-.02 + I +

```

20  -.09          XXI +
21  .05           + IX+
22  .07           + IXX
23  .00           + I +
24  -.05          +XI +

```

PARTIAL AUTOCORRELATIONS

```

1- 12   .00 -.05 -.04 -.04 .08 .09 -.08 -.03 .02 -.08 -.02 -.02
ST.E.   .04 .04 .04 .04 .04 .04 .04 .04 .04 .04 .04 .04

13- 24  -.07 .00 -.05 -.06 -.01 .02 -.02 -.11 .06 .07 -.03 -.06
ST.E.   .04 .04 .04 .04 .04 .04 .04 .04 .04 .04 .04 .04

```

```

      -1.0  -.8  -.6  -.4  -.2  .0  .2  .4  .6  .8  1.0
      +-----+-----+-----+-----+-----+-----+
              I
1   .00          + I +
2  -.05          +XI +
3  -.04          +XI +
4  -.04          +XI +
5   .08          + IXX
6   .09          + IXX
7  -.08          XXI +
8  -.03          +XI +
9   .02          + I +
10 -.08          XXI +
11 -.02          + I +
12 -.02          + I +
13 -.07          XXI +
14 .00           + I +
15 -.05          +XI +
16 -.06          +XI +
17 -.01          + I +
18 .02           + I +
19 -.02          + I +
20 -.11          X+XI +
21 .06           + IX+
22 .07           + IXX
23 -.03          +XI +
24 -.06          XXI +
--

```

IDENTIFY VARIABLE RESY.

```

NAME OF THE SERIES . . . . . RESY
TIME PERIOD ANALYZED . . . . . 11 TO 744
MEAN OF THE (DIFFERENCED) SERIES . . . . . .7338
STANDARD DEVIATION OF THE SERIES . . . . . 9.1869
T-VALUE OF MEAN (AGAINST ZERO) . . . . . 2.1639

```

AUTOCORRELATIONS

```

1- 12   .00 -.05 -.04 -.04 .08 .09 -.08 -.04 .01 -.07 .00 -.01
ST.E.   .04 .04 .04 .04 .04 .04 .04 .04 .04 .04 .04 .04
Q       .0  1.6  3.0  4.1  8.8 14.9 20.1 21.5 21.6 25.7 25.7 25.8

13- 24  -.08 .01 -.04 -.06 .01 .02 -.02 -.09 .05 .07 .00 -.05
ST.E.   .04 .04 .04 .04 .04 .04 .04 .04 .04 .04 .04 .04
Q      30.4 30.5 32.0 35.1 35.1 35.4 35.6 41.7 43.3 46.8 46.8 49.0

```

```

      -1.0  -.8  -.6  -.4  -.2  .0  .2  .4  .6  .8  1.0
      +-----+-----+-----+-----+-----+-----+
              I
1   .00          + I +
2  -.05          +XI +
3  -.04          +XI +
4  -.04          +XI +

```

```

5      .08      + IXX
6      .09      + IXX
7     -.08      XXI +
8     -.04      +XI +
9      .01      + I +
10    -.07      XXI +
11     .00      + I +
12    -.01      + I +
13    -.08      XXI +
14     .01      + I +
15    -.04      +XI +
16    -.06      XXI +
17     .01      + I +
18     .02      + IX+
19    -.02      + I +
20    -.09      XXI +
21     .05      + IX+
22     .07      + IXX
23     .00      + I +
24    -.05      +XI +

```

PARTIAL AUTOCORRELATIONS

```

1- 12      .00 -.05 -.04 -.04 .08 .09 -.08 -.03 .02 -.08 -.02 -.02
ST.E.      .04 .04 .04 .04 .04 .04 .04 .04 .04 .04 .04 .04

13- 24     -.07 .00 -.05 -.06 -.01 .02 -.02 -.11 .06 .07 -.03 -.06
ST.E.      .04 .04 .04 .04 .04 .04 .04 .04 .04 .04 .04 .04

```

```

-1.0  -.8  -.6  -.4  -.2  .0  .2  .4  .6  .8  1.0
+-----+-----+-----+-----+-----+-----+-----+-----+

```

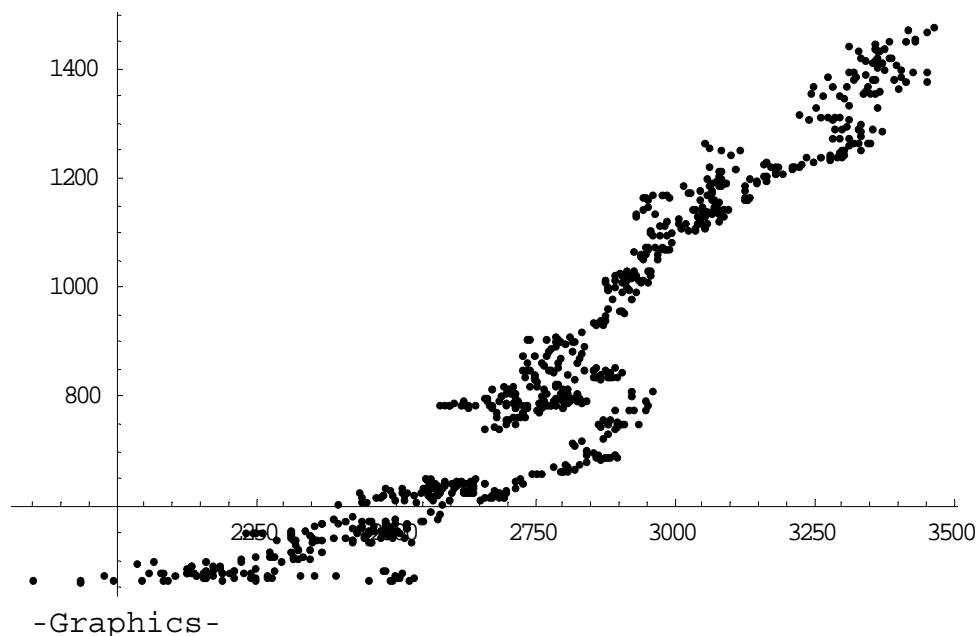
```

I
1      .00      + I +
2     -.05      +XI +
3     -.04      +XI +
4     -.04      +XI +
5      .08      + IXX
6      .09      + IXX
7     -.08      XXI +
8     -.03      +XI +
9      .02      + I +
10    -.08      XXI +
11    -.02      + I +
12    -.02      + I +
13    -.07      XXI +
14     .00      + I +
15    -.05      +XI +
16    -.06      +XI +
17    -.01      + I +
18     .02      + I +
19    -.02      + I +
20    -.11      X+XI +
21     .06      + IX+
22     .07      + IXX
23    -.03      +XI +
24    -.06      XXI +
--

```

## Výstup z programu MATHEMATICA

```
<<Graphics`MultipleListPlot`
<<Statistics`
<<Statistics`LinearRegression`
<<Statistics`ConfidenceIntervals`
<<Statistics`HypothesisTests`
SkalarniSoucin=Function[{u,v},
  Sum[u[[i]]*v[[i]],{i,1,Length[u]}]
];
ListPlot[data]
```



```
FSum[a_,b_,d_,j_]:=Sum[(data[[i,2]]-
If[i<=j,a+b*data[[i,1]],a+b*data[[IntegerPart[j],1]]+d*(data[[i,
1]]-data[[IntegerPart[j],1]])]^2,{i,1,Length[data]}];
General :: spell : Possible spelling error: new
symbol name "FSum" is similar to existing symbols FSun, Sum . More..

j=305;
g=Table[FindMinimum[FSum[a,b,d,j],{{a,0},{b,1},{d,1}}],{j,1,Le
ngth[data]}]

Min[h=Table[g[[k,1]],{k,1,Length[data]}]]
```



```

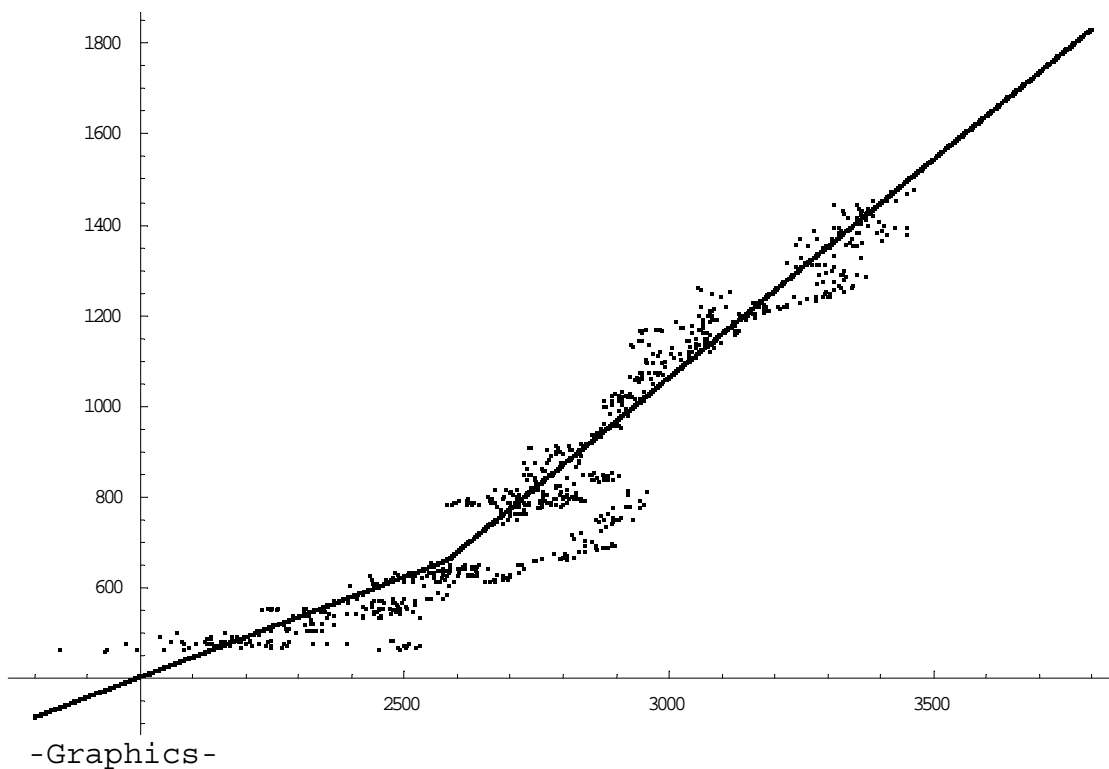
2.55056`106
Position[h,2.550557572235704`*6]
{{415}}
j=415;
vysl=FindMinimum[FSum[a,b,d,j],{{a,0},{b,1},{d,1}}]
2.55056`106, a⊗481.277, b⊗0.441951, d⊗0.961553
data[[415,1]]
2580.04
f[x_]:=If[x≤2580,-481.2771240512129+0.44195063585178496*x,-
481.2771240512129+0.44195063585178496*2580+0.9615530962628105*
(x-2580)];
fd=Table[{x,f[x]},{x,1800,3800}];

```

```

MultipleListPlot[data,fd,PlotJoined→{False,True},PlotRange→All,
SymbolShape→{PlotSymbol[Box,1],PlotSymbol[Box,1]}]

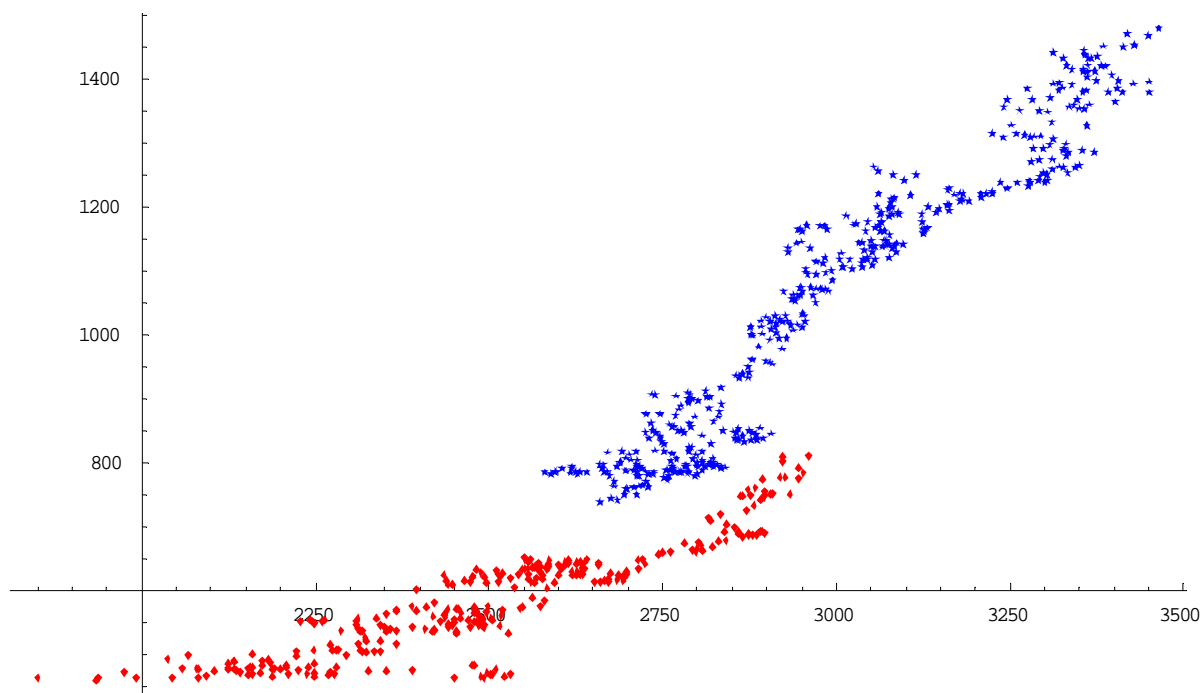
```



```

j=304;
data1=Table[{Euro[[i]],PX[[i]]},{i,1,j}];
data2=Table[{Euro[[i]],PX[[i]]},{i,j+1,Length[PX]}];
MultipleListPlot[data1,data2,SymbolStyle→{Red,Blue}]

```



-Graphics-

```

j=304;
pred=Table[{Euro[[i]],PX[[i]]},{i,1,j}];
po=Table[{Euro[[i]],PX[[i]]},{i,j+1,Length[PX]}];

```

### (\*Regresní model před\*)

```

(regresspred=Regress[pred,{1,x},x];
Chop[regresspred,10^(-6)])

```

ParameterTable®	Estimate	SE	TStat	PValue
1	-225.858	21.854	-10.3349	0
x	0.325	0.00872666	37.2422	0

RSquared® 0.821194, AdjustedRSquared® 0.820602,

EstimatedVariance® 1305.82, ANOVATable®

	DF	SumOfSq	MeanSq	FRatio	PValue
Model	1	$1.81115 \cdot 10^6$	$1.81115 \cdot 10^6$	1386.98	0
Error	302	394359.	1305.82		
Total	303	$2.20551 \cdot 10^6$			

### (\*Regresní model po\*)

```

(regresspo=Regress[po,{1,x},x];
Chop[regresspo,10^(-6)])

```

	Estimate	SE	TStat	PValue
ParameterTable <sup>®</sup> 1	-1644.3	36.832	-44.6431	0
x	0.902417	0.0122879	73.4392	0

RSquared<sup>®</sup> 0.924888, AdjustedRSquared<sup>®</sup> 0.924717,  
 EstimatedVariance<sup>®</sup> 3495.47, ANOVATable<sup>®</sup>

	DF	SumOfSq	MeanSq	FRatio	PValue
Model	1	1.88522 · 10 <sup>7</sup>	1.88522 · 10 <sup>7</sup>	5393.32	0
Error	438	1.53101 · 10 <sup>6</sup>	3495.47		
Total	439	2.03832 · 10 <sup>7</sup>			



```
X1=Table[{1,pred[[i,1]]},{i,1,Length[pred]}];
X2=Table[{1,po[[i,1]]},{i,1,Length[po]}];
Y1=Map[Last,pred];
Y2=Map[Last,po];
b1=Inverse[Transpose[X1].X1].Transpose[X1].Y1;
b2=Inverse[Transpose[X2].X2].Transpose[X2].Y2;
Se1=SkalarniSoucin[(Y1-X1.b1),(Y1-X1.b1)];
Se2=SkalarniSoucin[(Y2-X2.b2),(Y2-X2.b2)];
v111=Inverse[Transpose[X1].X1][[2,2]];
v112=Inverse[Transpose[X2].X2][[2,2]];
```

```
T=(b1[[2]]-b2[[2]])*Sqrt[Length[pred]+Length[po]-4]/(((Se1+Se2)*(v111+v112))^(1/2))
```

-35.5288

```
Q=Quantile[StudentTDistribution[Length[pred]+Length[po]-4],0.95]
1.64692
```

## (\*Výsledek testu na hladině spolehlivosti 95%\*)

```
If[Abs[T]>=Q,Print["Zamítáme hypotézu stejné sm•rnice regresní p•ímky p•ed Temelínem a po Temelínu"],Print["Nezamítáme hypotézu stejné sm•rnice regresní p•ímky p•ed Temelínem a po Temelínu"]]
Zamítáme hypotézu stejné sm•rnice regresní p•ímky p•ed Temelínem a po Temelínu
```

```
Variance[PX]*(Length[PX]-1)
6.20204 · 107
```

```
SKvadratLomenaRegrese=1-2.550557572235704`*^6/6.202040059124664`*^7
```

0.958876

```
SKvadratdveregrese=1-
(394358.8285801744+1531014.4659121593)/6.202040059124664`*^7
0.968956
```

Tabulka 1 – Jednodenní předpovědi volatility míry zisku indexu PX 50

Datum	t	Index PX 50	$r_t(\%)$	$r_t^2$	$\sigma_{t+1 t}^2$	$\sigma_{t+1 t}$
11.9.2001	0	340,0	-1,762	3,106	3,106	1,762
12.9.2001	1	330,5	-2,794	7,807	3,388	1,841
13.9.2001	2	335,4	1,483	2,198	3,317	1,821
14.9.2001	3	322,9	-3,727	13,890	3,951	1,988
17.9.2001	4	320,1	-0,867	0,752	3,759	1,939
18.9.2001	5	321,2	0,344	0,118	3,541	1,882
19.9.2001	6	330,9	3,020	9,120	3,876	1,969
20.9.2001	7	327,8	-0,937	0,878	3,696	1,922
21.9.2001	8	325,2	-0,793	0,629	3,512	1,874
24.9.2001	9	328,0	0,861	0,741	3,346	1,829
25.9.2001	10	332,7	1,433	2,053	3,268	1,808
26.9.2001	11	337,0	1,292	1,670	3,172	1,781
27.9.2001	12	331,9	-1,513	2,290	3,119	1,766
1.10.2001	13	333,9	0,603	0,363	2,954	1,719
2.10.2001	14	339,6	1,707	2,914	2,952	1,718
3.10.2001	15	334,7	-1,443	2,082	2,899	1,703
4.10.2001	16	340,4	1,703	2,900	2,899	1,703
5.10.2001	17	340,3	-0,029	0,001	2,725	1,651
8.10.2001	18	343,1	0,823	0,677	2,603	1,613
9.10.2001	19	344,2	0,321	0,103	2,453	1,566
10.10.2001	20	350,6	1,859	3,457	2,513	1,585
11.10.2001	21	359,2	2,453	6,017	2,723	1,650
12.10.2001	22	354,6	-1,281	1,640	2,658	1,630
15.10.2001	23	351,3	-0,931	0,866	2,551	1,597
16.10.2001	24	352,6	0,370	0,137	2,406	1,551
17.10.2001	25	359,1	1,843	3,398	2,465	1,570
18.10.2001	26	359,4	0,084	0,007	2,318	1,522
19.10.2001	27	361,9	0,696	0,484	2,208	1,486
22.10.2001	28	363,6	0,470	0,221	2,089	1,445

23.10.2001	29	364,8	0,330	0,109	1,970	1,403
24.10.2001	30	368,0	0,877	0,769	1,898	1,378
25.10.2001	31	368,6	0,163	0,027	1,785	1,336
26.10.2001	32	376,6	2,170	4,711	1,961	1,400
29.10.2001	33	376,8	0,053	0,003	1,844	1,358
30.10.2001	34	372,5	-1,141	1,302	1,811	1,346
31.10.2001	35	375,3	0,752	0,565	1,736	1,318
1.11.2001	36	378,6	0,879	0,773	1,678	1,296
2.11.2001	37	381,4	0,740	0,547	1,611	1,269
5.11.2001	38	382,3	0,236	0,056	1,517	1,232
6.11.2001	39	384,5	0,575	0,331	1,446	1,203
7.11.2001	40	386,5	0,520	0,271	1,376	1,173
8.11.2001	41	391,0	1,164	1,356	1,374	1,172
9.11.2001	42	387,5	-0,895	0,801	1,340	1,158
12.11.2001	43	383,3	-1,084	1,175	1,330	1,153
13.11.2001	44	392,4	2,374	5,636	1,588	1,260
14.11.2001	45	399,7	1,860	3,461	1,701	1,304
15.11.2001	46	398,0	-0,425	0,181	1,610	1,269
16.11.2001	47	401,4	0,854	0,730	1,557	1,248
19.11.2001	48	409,1	1,918	3,680	1,684	1,298
20.11.2001	49	405,3	-0,929	0,863	1,635	1,279
21.11.2001	50	403,0	-0,567	0,322	1,556	1,247
22.11.2001	51	407,7	1,166	1,360	1,544	1,243
23.11.2001	52	408,5	0,196	0,039	1,454	1,206
26.11.2001	53	407,5	-0,245	0,060	1,370	1,171
27.11.2001	54	398,2	-2,282	5,208	1,601	1,265
28.11.2001	55	393,4	-1,205	1,453	1,592	1,262
29.11.2001	56	393,6	0,051	0,003	1,496	1,223
30.11.2001	57	392,5	-0,279	0,078	1,411	1,188
3.12.2001	58	395,1	0,662	0,439	1,353	1,163
4.12.2001	59	402,1	1,772	3,139	1,460	1,208
5.12.2001	60	401,3	-0,199	0,040	1,375	1,173

6.12.2001	61	401,0	-0,075	0,006	1,293	1,137
7.12.2001	62	396,4	-1,147	1,316	1,294	1,138
10.12.2001	63	393,3	-0,782	0,612	1,253	1,119
11.12.2001	64	392,9	-0,102	0,010	1,179	1,086
12.12.2001	65	394,7	0,458	0,210	1,121	1,059
13.12.2001	66	392,8	-0,481	0,232	1,067	1,033
14.12.2001	67	390,1	-0,687	0,472	1,031	1,016
17.12.2001	68	383,6	-1,666	2,776	1,136	1,066
18.12.2001	69	384,3	0,182	0,033	1,070	1,034
19.12.2001	70	396,1	3,071	9,428	1,572	1,254
20.12.2001	71	392,9	-0,808	0,653	1,516	1,231
21.12.2001	72	391,3	-0,407	0,166	1,435	1,198
27.12.2001	73	393,3	0,511	0,261	1,365	1,168
28.12.2001	74	394,6	0,331	0,109	1,290	1,136

Tabulka 2 – Jednodenní předpovědi volatility míry zisku indexu Standard & Poor's 500

Datum	t	Index Standard & Poor's 500	$r_t(\%)$	$r_t^2$	$\sigma_{t+1 t}^2$	$\sigma_{t+1 t}$
10.9.2001	0	1 092,54	0,623	0,388	0,388	0,623
17.9.2001	1	1 038,77	-4,922	24,222	1,818	1,348
18.9.2001	2	1 032,74	-0,580	0,337	1,729	1,315
19.9.2001	3	1 016,10	-1,611	2,596	1,781	1,334
20.9.2001	4	984,54	-3,106	9,647	2,253	1,501
21.9.2001	5	965,80	-1,903	3,623	2,335	1,528
24.9.2001	6	1 003,45	3,898	15,197	3,107	1,763
25.9.2001	7	1 012,27	0,879	0,773	2,967	1,722
26.9.2001	8	1 007,04	-0,517	0,267	2,805	1,675
27.9.2001	9	1 018,61	1,149	1,320	2,716	1,648
28.9.2001	10	1 040,94	2,192	4,806	2,841	1,686
1.10.2001	11	1 038,55	-0,230	0,053	2,674	1,635
2.10.2001	12	1 051,33	1,231	1,514	2,604	1,614
3.10.2001	13	1 072,28	1,993	3,971	2,686	1,639
4.10.2001	14	1 069,62	-0,248	0,062	2,529	1,590
5.10.2001	15	1 071,38	0,165	0,027	2,379	1,542
8.10.2001	16	1 062,44	-0,834	0,696	2,278	1,509
9.10.2001	17	1 056,75	-0,536	0,287	2,158	1,469
10.10.2001	18	1 080,99	2,294	5,262	2,344	1,531
11.10.2001	19	1 097,43	1,521	2,313	2,343	1,531
12.10.2001	20	1 091,65	-0,527	0,277	2,219	1,490
15.10.2001	21	1 089,98	-0,153	0,023	2,087	1,445
16.10.2001	22	1 097,54	0,694	0,481	1,991	1,411
17.10.2001	23	1 077,09	-1,863	3,472	2,079	1,442
18.10.2001	24	1 068,61	-0,787	0,620	1,992	1,411
19.10.2001	25	1 073,48	0,456	0,208	1,885	1,373

22.10.2001	26	1 089,90	1,530	2,340	1,912	1,383
23.10.2001	27	1 084,78	-0,470	0,221	1,811	1,346
24.10.2001	28	1 085,20	0,039	0,001	1,702	1,305
25.10.2001	29	1 100,09	1,372	1,883	1,713	1,309
26.10.2001	30	1 104,61	0,411	0,169	1,620	1,273
29.10.2001	31	1 078,30	-2,382	5,673	1,863	1,365
30.10.2001	32	1 059,79	-1,717	2,947	1,928	1,389
31.10.2001	33	1 059,78	-0,001	0,000	1,813	1,346
1.11.2001	34	1 084,10	2,295	5,266	2,020	1,421
2.11.2001	35	1 087,20	0,286	0,082	1,904	1,380
5.11.2001	36	1 102,84	1,439	2,069	1,914	1,383
6.11.2001	37	1 118,86	1,453	2,110	1,925	1,388
7.11.2001	38	1 115,80	-0,273	0,075	1,814	1,347
8.11.2001	39	1 118,54	0,246	0,060	1,709	1,307
9.11.2001	40	1 120,31	0,158	0,025	1,608	1,268
12.11.2001	41	1 118,33	-0,177	0,031	1,513	1,230
13.11.2001	42	1 139,09	1,856	3,446	1,629	1,276
14.11.2001	43	1 141,21	0,186	0,035	1,534	1,238
15.11.2001	44	1 142,24	0,090	0,008	1,442	1,201
16.11.2001	45	1 138,65	-0,314	0,099	1,362	1,167
19.11.2001	46	1 151,06	1,090	1,188	1,351	1,162
20.11.2001	47	1 142,66	-0,730	0,533	1,302	1,141
21.11.2001	48	1 137,03	-0,493	0,243	1,238	1,113
23.11.2001	49	1 150,34	1,171	1,370	1,246	1,116
26.11.2001	50	1 157,42	0,615	0,379	1,194	1,093
27.11.2001	51	1 149,50	-0,684	0,468	1,151	1,073
28.11.2001	52	1 128,52	-1,825	3,331	1,282	1,132
29.11.2001	53	1 140,20	1,035	1,071	1,269	1,126
30.11.2001	54	1 139,45	-0,066	0,004	1,193	1,092
3.12.2001	55	1 129,90	-0,838	0,702	1,164	1,079
4.12.2001	56	1 144,80	1,319	1,739	1,198	1,095



5.12.2001	57	1 170,35	2,232	4,981	1,425	1,194
6.12.2001	58	1 167,10	-0,278	0,077	1,344	1,159
7.12.2001	59	1 158,31	-0,753	0,567	1,298	1,139
10.12.2001	60	1 139,93	-1,587	2,518	1,371	1,171
11.12.2001	61	1 136,76	-0,278	0,077	1,293	1,137
12.12.2001	62	1 137,07	0,027	0,001	1,216	1,103
13.12.2001	63	1 119,38	-1,556	2,420	1,288	1,135
14.12.2001	64	1 123,09	0,331	0,110	1,217	1,103
17.12.2001	65	1 134,36	1,003	1,007	1,205	1,098
18.12.2001	66	1 142,92	0,755	0,569	1,167	1,080
19.12.2001	67	1 149,56	0,581	0,338	1,117	1,057
20.12.2001	68	1 139,93	-0,838	0,702	1,092	1,045
21.12.2001	69	1 144,89	0,435	0,189	1,038	1,019
24.12.2001	70	1 144,65	-0,021	0,000	0,976	0,988
26.12.2001	71	1 149,37	0,412	0,170	0,927	0,963
27.12.2001	72	1 157,13	0,675	0,456	0,899	0,948
28.12.2001	73	1 161,02	0,336	0,113	0,852	0,923
31.12.2001	74	1 148,08	-1,115	1,242	0,875	0,936

## Literatura

- [1] Abraham, B. - Ledolter, J. (1983): Statistical Methods for Forecasting, John Wiley & Sons, New York.
- [2] Anderson, O. D.(1976): Time series analysis and forecasting - Box-Jenkins approach. London, Butterworth .
- [3] Anderson, O.D. (1976): On the transformation of raw time series data-a review, Statistische Hefte, 17, 285-289.
- [4] Anděl , J (1978) : Matematická statistika, SNTL/ALFA, Praha..
- [5] Anděl, J. (1998): Statistické metody, MATFYZPRESS, Praha.
- [6] Arlt, J., Arltová, M. (2003): Finanční časové řady, Grada Publishing, Praha.
- [7] Baillie,R.T.- Bollerslev,T.-Mikkelsen,H.O.(1996): Fractionally Integrated Generalized Autoregressive Conditional Heteroscedasticity, Journal of Econometrics, 74, 3-30.
- [8] Barnard, G.A.-Jenkins, G.M.-Winstern, C.B. (1962): Likelihood inference and time series, J.Roy, Statist.Soc., A125, 321-352.
- [9] Bartlett, M.S.(1946): On the theoretical specification of sampling properties of autocorrelated time series, J.Roy.Statist.Soc., B8, 27-41
- [10] Berben, R.-P.- van Dijk, D.(1999): Unit Root Tests and Asymmetric Adjustment – a Reassessment, Econometric Institute Report 9902, Erasmus University Rotterdam.
- [11] Beran,J (1995).: Maximum Likelihood Estimation of the Differencing Parameter for Invertible Short and Long Memory Autoregressive Integrated Moving Average Models, Journal of the Royal Statistical Society B57, 654-672.
- [12] Bollerslev, T.(1986): Generalized Autoregressive Conditional Heteroskedasticity, Journal of Econometrics, 31, 307-327.
- [13] Box, G.E.P.-Cox, D.R.(1964): An analysis of transformstions, J.Roy.Statist.Soc., B26, 211-252.
- [14] Box, G.E.P.-Jenkins, G.M.(1970): Time series analysis, forecasting and control, Holden Day, San Francisco.

- [15] Box, G. E. P. - Jenkins, G. M. - Reinsel, G.C.(1994): Time series analysis, forecasting and control. Third edition. Prentice-Hall, Inc. Englewood Cliffs, New Jersey.
- [16] Box, G.E.P. – Tiao, G.C.(1975): Intervention analysis with applications to economic and environmental problems, *JASA* 70, 70-79.
- [17] Brockwell, P.J.-Davis, R.A. (1996): Time Series: Theory and Methods, (2<sup>nd</sup> ed.), Springer-Verlag, New York..
- [18] Brown, B.Y. – Mariano, R.S.(1989): Predictors in Dynamic Nonlinear Models : Large Sample Behaviour, *Econometric Theory* 5, 430-452.
- [19] Caner,M. – Hansen, B.E.(1997): Threshold Autoregressions with a Unit Root, Working Papers in Economics 381, Boston College.
- [20] Cipra, T. (1986): Analýza časových řad s aplikacemi v ekonomii, SNTL, Praha.
- [21] Cipra, T.(2002): Kapitálová přiměřenost ve financích a solventnost v pojišťovnictví, Ekopress , Praha.
- [22] de Gooijer, J.G.- de Bruin, P (1998).: On Forecasting SETAR Processes, *Statistics and Probability Letters* 37, 7-14.
- [23] Enders, W. – Granger, C.W.J(1998).: Unit-Root Tests and Asymmetric Adjustment with an Example Using the Term Structure of Interest Rates, *Journal of Business & Economic Statistics* 16, 304-311.
- [24] Engle, R.F.(1982): Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica* 50, 987-1007.
- [25] Engle, R.F.-Lilien,D.M. – Robins,R.P.(1987): Estimating Time Varying Risk Premia in the Term Structure. The ARCH-M Model, *Econometrica* 55, 391-407.
- [26] Engle, R.F.-Ng,V.K.(1993): Measuring and Testing the Impact of News on Volatility, *Journal of Finance* 48, 1749 – 1778.
- [27] Fichtengolc, G.M.(1962): Kurs diferencialnogo i integralnogo isčislenija II, Moskva.
- [28] Fornari,F.-Mele,A.(1996): Modeling the Changing Asymmetry of Conditional Variances, *Economics Letters* 50, 197-203.
- [29] Frances, P.H.-van Dijk, D.(2000): Non-linear Time Series Models in Empirical Finance, Cambridge University Press, Cambridge.

- 
- [30] Frances, P.H. – Ghijssels,H.(1999): Additive Outliers, GARCH and Forecasting Volatility, *International Journal of Forecasting* 15, 1-9.
- [31] *Forecasting and Time Series Analysis using the SCA Statistical System. Volume 1.* Scientific Associates Corp. Oak Brook, Illinois, USA, 1991.
- [32] Geweke,J.-Porter-Hudak,S.(1983): The Estimation and Application of Long Memory Time Series Models, *Journal of Time Series Analysis*, 4, 221-238.
- [33] Glosten,L.R.-Jagannathan, R.-Runkle,D.E.(1993): On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stock, *Journal of Finance* 48, 1779-1801.
- [34] Gonzales-Rivera,G.(1998): Smooth Transition GARCH Models, *Studies in Nonlinear Dynamics and Econometrics* 3, 61-78.
- [35] Granger, C.W.J.(1980): Long Memory Relationships and the Aggregation of Dynamic Models, *Journal of Econometrics* 14, 227-238.
- [36] Granger, C.W.J.-Joyeux, R(1980).: An Introduction to Long Memory Time Series Models and Fractional Differencing, *Journal of Time Series Analysis* 1, 15-29.
- [37] Granger, C. W. J., Newbold , P.(1986): *Forecasting Economic Time Series.* Academic Press. New York.
- [38] Hagerud,G.E.(1997): A New Non-Linear GARCH Model, PhD Thesis, IFE, Stockholm School of Economics.
- [39] Hamilton,J.D.(1989): A New Approach to the Economic Analysis of Nonstationary Time Series Subject to Changes in Regime, *Econometrica* 57, 357-384.
- [40] Hamilton, J.D.(1990): Analysis of Time series Subject to Changes in Regime, *Journal of Econometrics* 45, 39-70.
- [41] Hamilton J.D. (1993): Estimation, Inference and Forecasting of Time Series Subject to Changes in Regime, in Maddala. G.S., Rao, C.R., Vinod, H.D.(eds.), *Handbook of Statistics* 11, Amsterdam, North-Holland, 231-260.
- [42] Hamilton, J.D. (1994): *Time Series Analysis*, Princeton University Press, Princeton.
- [43] Hamilton, J.D.(1996) : Specification Testing in Markov-Switching Time Series Models, *Journal of Applied Econometrics* 70, 127-157.

- 
- [44] Hansen, B.E.(1992): The Likelihood Ratio Test Under Nonstandard Assumptions: Testing the Markov Switching Models GNP, *Journal of Applied Econometrics* 7, S61-S82; erratum11, 195-198.
- [45] Hansen, B.E.(1997): Inference in TAR models, *Studies in Nonlinear Dynamics and Econometrics* 2, 1-14.
- [46] Hansen, B.E.(2000): Sample Splitting and Threshold Estimation, *Econometrica* 68, Issue 3, 575-603.
- [47] Harvey, A.C. (1981): *The econometric analysis of time series*, Oxford, Allan.
- [48] Hillmer, S.C. – Tiao, G.C.(1979): Likelihood Function of Stationary Multiple Autoregressive – Moving Average Models, *JASA* 74, 652-660.
- [49] Hosking, J.R.M. (1981): Equivalent Forms of the Multivariate Portmanteau Statistic, *Journal of the Royal Statistical Society B* 43, 261-262.
- [50] Hurst, H.(1951): Long Term Storage Capacity of Reservoirs, *Transactions of the American Society of Civil Engineers* 116, 770-799.
- [51] Chan, K.S.(1990): Testing for Threshold Autoregression, *Annals of Statistics* 18, 1886-1894.
- [52] Chan, K.S.(1991): Percentage Points for Likelihood Ratio Tests for Threshold Autoregression, *Journal of the Royal Statistical Society B* 53, 691-696.
- [53] Chan, K.S.-Tong,H.(1990): On Likelihood Ratio Tests for Threshold Autoregression, *Journal of the Royal Statistical Society B* 52, 469-476.
- [54] Chatfield, C.-Prothero, D.L.(1973): Box-Jenkins seasonal forecasting – problems in a case study, *J.Roy.Statist.Soc., A* 136, 295-314.
- [55] Chen, C.-Liu, L.M.(1993): Joint Estimation of Model Parameters and Outlier Effects in Time Series, *Journal of the American Statistical Association* 88, 284-297.
- [56] Koreisha, S.G.-Taylor, S.A.(1985): Identification of Transfer Function Models: An Asymptotic Test of Significance for the corner Method, *Communications in Statistics, Part A – Theory and Methods* 14, 159 – 173.
- [57] Jarník, V.(1976): *Diferenciální počet*, ACADEMIA, Praha.
- [58] Liu, L.M. (1984): Estimation of Rational Transfer Function Models, *Communications in Statistics – Simulationand Computation* 13, 775-784.

- 
- [59] Ljung, G.M.- Box, G.E.P.(1978): On a Measure of Lack of Fit in Time Series Models, *Biometrika* 65, 297-303.
- [60] Lundbergh, S.-Teräsvirta, T.(1998b): Evaluating GARCH Models, Working Papers in Economics and Finance 292, Stockholm School of Economic.
- [61] Luukkonen, R.-Saikkonen,P.-Teräsvirta, T.(1988): Testing Linearity Against Smooth Transition Autoregressive Models, *Biometrika* 75, 49-497.
- [62] Marek,L.(1993): Intervenční analýza v časových řadách, *Statistika* 30, 429-438.
- [63] Marek,L.(1996): Odlehlá pozorování v časových řadách, *Acta economica pragensia*, vědecký sborník Vysoké školy ekonomické v Praze: Statistické modelování v hospodářské praxi, Praha.
- [64] Marek,L.(1998): Intervention Analysis and Outliers in Time Series, *Štatistické metódy v praxi*, Bratislava.
- [65] Marek,L.(1999): Building of Transfer Function Models, *Výpočtová štatistika*, Bratislava. Marek,L.(2000): Time Series Analysis in the SAS System, *Výpočtová štatistika*, Bratislava.
- [66] Marek,L.(2001): Modely s přenosovou funkcí, *Habilitační práce*, Praha.
- [67] Nelson,D.B.(1991): Conditional Heteroskedasticity in Asset Returns: a New Approach, *Econometrica* 59, 347-370.
- [68] Pankratz, A. (1991): *Forecasting with dynamic regression models*. John Wiley & Sons inc., New York.
- [69] Quenouille, M.H.(1949): Approximate tests of correlation in time series, *J.Roy.Statist.Soc.*, B11, 68-84.
- [70] Sekerka, B (1996): *Cenné papíry a kapitálové trhy*, Profess, Praha.
- [71] Sentana, E.(1995): Quadratic ARCH Models, *Review of Economic Studies* 62, 639-661.
- [72] Sowell, F.B.(1986): *Fractionally Integrated Vector Time Series*, PhD.Dissertation, Duke University, Durham.
- [73] Sowel, F.B.(1992): Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models, *Journal of Econometrics* 53, 165-188.
- [74] The SCA Statistical System. Reference manual for general statistical analysis. Scientific Associates Corp. Oak Brook, Illinois, USA, 1991.

- 
- [75] StatSoft, Inc. (2001). STATISTICA (data analysis software system), version 6. [www.statsoft.com](http://www.statsoft.com).
- [76] Taylor,S.(1986): Modelling Financial Time Series, Wiley, Chichester.
- [77] Teräsvirta, T.(1994): Specification, Estimation and Evaluation of Smooth Transition Autoregressive Models, Journal of the American Statistical Association 89, 208-218.
- [78] Teräsvirta, T.(1996): Two Stylized Fact and the GARCH(1,1) Model, Working Paper Series in Economics and Finance 96, Stockholm School of Economics.
- [79] Teräsvirta, T.(1998): Modelling Economic Relationships with Smooth Transition Regressions, in Ullah,A. and Giles, D.E.A. (eds.), Handbook of Applied Economic Statistics, Marcel Dekker, New York, 507-552.
- [80] Tong, H.(1978): On a Threshold Model in Chen, C.H. (eds.), Pattern Recognition and Signal Processing, Sijhoff & Noordhoff, Amsterdam, 101-141.
- [81] Tong,H.(1990): Non-Linear Time Series: A Dynamical Systems Approach, Oxford University Press, Oxford.
- [82] Tong, H. – Lim, K.S.(1980): Threshold Autoregressions , Limit Cycles, and Data, Journal of the Royal Statistical Society B 42, 245 –292.
- [83] Trešl, J. (1999): Statistické metody a kapitálové trhy, Skripta VŠE, Praha.
- [84] Trešl, J.(2005): Statistical analysis of PX –50 time series, Aplimat 2005, Bratislava, 549-554.
- [85] Wall, K.D.(1976): FIML Estimation of Rational Distributed Lag Structural Form Models, Annals of Economic and Social Measurement 5, 53-64.
- [86] Wolfram Research, Inc. (2005). Mathematica (mathematical software system), version 5.2. [www.wolfram.com](http://www.wolfram.com).
- [86] Wei, W. W. S.(1990): Time series analysis - Univariate and multivariate methods. Redwood City, California, Addison-Wesley Publishing Company .