

Česká zemědělská univerzita v Praze

Provozně ekonomická fakulta

Katedra statistiky



Diplomová práce

**Návrh a implementace obchodního systému
pro obchodování komodit**

Bc. Jan Voldán

© 2019 ČZU v Praze

ZADÁNÍ DIPLOMOVÉ PRÁCE

Bc. Jan Voldán

Provoz a ekonomika

Název práce

Návrh a implementace obchodního systému pro obchodování komodit

Název anglicky

Design and implementation of a trading system for trading commodities

Cíle práce

Hlavním cílem práce je vytvořit obchodní systém, který bude obchodovat s vybranými komoditami, za využití matematicko-statistických modelů ARIMA a GARCH.

Díličními cíli jsou:

- vytvoření kvalitních predikcí ceny vybraných komodit na základě ARIMA/GARCH modelů v R
- verifikace predikcí formou backtestu v prostředí R
- implementace obchodního systému do prostředí MetaTrader4
- ověření vytvořeného systému na reálných datech.

Metodika

Těžiště diplomové práce spočívá v tvorbě obchodního systému, který bude programován v jazyce MQL4 na platformě MetaTrader4, jehož hlavní výpočty budou programovány v jazyce R. K tvorbě systému budou využity modely GARCH, ARIMA. Funkčnost systému bude dále ověřována na reálných datech po dobu jednoho měsíce v prostředí MetaTrader4.

Doporučený rozsah práce

60 – 80 stran

Klíčová slova

Obchodní systém, GARCH, ARIMA, komodita

Doporučené zdroje informací

ARLTOVÁ, M. – ARLT, J. *Ekonomické časové řady : [vlastnosti, metody modelování, příklady a aplikace]*.

Praha: Grada, 2007. ISBN 978-80-247-1319-9.

CIPRA, T. *Finanční ekonometrie*. Praha: Ekopress, 2013. ISBN 978-80-86929-93-4.

FIO BANKA, – REJNUŠ, O. *Finanční trhy*. Praha: Grada, 2014. ISBN 978-80-247-3671-6.

JANÁČ, R. – Automatické obchodní systémy, aneb, Forex Robot. Brno: Tribun EU, 2016. ISBN 978-80-263-1055-6.

JÍLEK, J. *Finanční a komoditní deriváty v praxi*. Praha: Grada, 2010. ISBN 978-80-247-3696-9.

NESNÍDAL, T. – PODHAJSKÝ, P. *Obchodování na komoditních trzích : průvodce spekulanta*. Praha: Grada, 2006. ISBN 80-247-1851-0.

Předběžný termín obhajoby

2019/20 ZS – PEF (únor 2020)

Vedoucí práce

Ing. Tomáš Hlavsa, Ph.D.

Garantující pracoviště

Katedra statistiky

Elektronicky schváleno dne 15. 1. 2019

prof. Ing. Libuše Svatošová, CSc.

Vedoucí katedry

Elektronicky schváleno dne 5. 2. 2019

Ing. Martin Pelikán, Ph.D.

Děkan

V Praze dne 28. 11. 2019

Čestné prohlášení

Prohlašuji, že svou diplomovou práci "Návrh a implementace obchodního systému pro obchodování komodit" jsem vypracoval samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou citovány v práci a uvedeny v seznamu použitých zdrojů na konci práce. Jako autor uvedené diplomové práce dále prohlašuji, že jsem v souvislosti s jejím vytvořením neporušil autorská práva třetích osob.

V Praze dne 29.11.2019

Poděkování

Jako první bych rád touto cestou poděkoval svému vedoucímu diplomové práce panu Ing. Tomáši Hlavsovi, Ph.D. za uvedení do problematiky ARIMA/GARCH modelů, jeho ochotu, velmi cenné náměty při psaní této práce a za podnětné připomínky. Dále bych chtěl poděkovat svému manažerovi ve své profesi, panu Ing. Lukášovi Petercovi za volnost, kterou mi poskytl na vypracovávání diplomové práce. Velké díky patří i mé rodině, která mě podporovala nejen při psaní diplomové práce, ale i během celého studia.

Návrh a implementace obchodního systému pro obchodování komodit

Abstrakt

Diplomová práce se zabývá problematikou obchodního systému pro obchodování s komoditami. Obchodní systém obsažený v této práci využívá smíšený model autoregresního integrovaného klouzavého průměru (ARIMA) a generalizovaného autoregresního podmíněného heteroskedastického modelu (GARCH). Nejprve je práce zpracovávána na teoretické úrovni a následně se práce zabývá samotnou problematikou návrhu, implementace a testování obchodního systému, jehož vytvoření byl stěžejní cíl této práce. Z výsledků testování bylo zjištěno, že systém by mohl být použitelný z dlouhodobého hlediska, jelikož provedené backtesty vytvořené v programu R na více než sedmiletém období vyšly pozitivně. Po implementaci R kódu do platformy Meta Trader 4 proběhlo otestování na reálných datech po dobu jednoho měsíce. Obchodování na reálných datech se ukázalo být naprosto nevhodné pro platinu z důvodu jejího vysokého spreadu u daného brokera. Obchodování na zlatě a ropě sice skončilo v profitu, avšak strategie vykazuje velkou volatilitu, a je proto pro krátkodobé obchodování nevhodná. Testování na reálných datech a vedení obchodního deníku zároveň přineslo i poznatky k vylepšení obchodního systému. Jde zejména o optimalizaci rychlosti a profitu na jeden obchod.

Klíčová slova: Obchodní systém, GARCH, ARIMA, komodita

Design and implementation of a trading system for trading commodities

Abstract

The thesis deals with the issue of trading system for commodities trading. The trading system included in this work uses a mixed model of the autoregressive integrated moving average (ARIMA) and the generalized autoregressive conditional heteroskedastic model (GARCH). First, the work is processed at the theoretical level and then the thesis deals with the issue of design, implementation and testing of the trading system, whose creation was the main objective of this work. From the results of the testing it was found that the system could be usable in the long term, because the backtests created in the R program for more than seven years were positive. After the implementation of the R code into the Meta Trader 4 platform, the real data was tested for one month. Trading on real data turned out to be unsuitable for platinum because of its high spread with the given broker. Trading in gold and oil ended in profit, but the strategy is highly volatile and is therefore unsuitable for short-term trading. At the same time, testing on real data and keeping a trading diary brought knowledge to improve the trading system. It is about optimizing speed and profit per trade.

Keywords: Trading system, GARCH, ARIMA, commodity

Obsah

1. Úvod	12
2. Cíl práce a metodika	14
2.1 Cíl práce	14
2.2 Metodika	14
2.2.1 Metodika výpočtů ARIMA/GARCH modelů	14
3. Teoretická východiska burzovního obchodování	31
3.1 Komodity	31
3.1.1 Členění komodit	31
3.1.2 Ropa	31
3.1.3 Platina	34
3.1.4 Zlato	36
3.2 Komoditní burza	39
3.3 Analýza	40
3.3.1 Psychologická analýza	40
3.3.2 Fundamentální analýza	42
3.3.3 Technická analýza	42
3.4 Indikátory technické analýzy	43
3.4.1 Indikátory	44
3.4.2 Oscilátory	45
3.4.3 Prediktivní indikátory	46
3.5 Risk management	46
3.5.1 Risk-Reward-Ratio (RRR)	47
3.6 Obchodní systémy	47
3.6.1 Typy obchodních platforem	49
3.6.2 MetaTrader 4	50
3.6.3 Rizika online obchodování	50
3.6.4 Backtest	51
4. Vlastní práce	53
4.1 Analýza v R	53
4.1.1 Použité funkce v programu R	53
4.1.2 Analýza a predikce v R	55
4.2 Obchodní systém	63
4.2.1 Backtest	63
4.2.2 Jazyk MQL a způsob použití	67
4.2.3 Aplikace na reálných datech	67

4.2.4	Pravidla obchodního systému	69
5.	Výsledky obchodování na reálných datech.....	71
5.1	Analýza výsledků strategie	71
5.2	Analýza jednotlivých komodit	73
5.2.1	Zlato	73
5.2.2	Ropa	75
5.2.3	Platina	77
5.3	Analýza obchodního deníku.....	79
6.	Závěr	84
6.1	Doporučení pro zlepšení	85
7.	Seznam použitých zdrojů	86
8.	Přílohy	89

Seznam obrázků

Obrázek 1	- Screen z programu Quant Analyzer 4 podrobné analýzy výsledků obchodování	71
Obrázek 2	- Screen z programu Quant Analyzer 4 podrobné analýzy výsledků obchodování zlata	73
Obrázek 3	- Screen z programu Quant Analyzer 4 podrobné analýzy výsledků obchodování ropy	75
Obrázek 4	- Screen z programu Quant Analyzer 4 podrobné analýzy výsledků obchodování platiny	77

Seznam tabulek

Tabulka 1	– Produkce ropy na den	32
Tabulka 2	– Množství ověřených ropných rezerv	33
Tabulka 3	- Porovnání jednotlivých modelů ARMA(p,q) dle AIC.....	58
Tabulka 4	- Odhadnuté parametry modelu ARIMA.....	59
Tabulka 5	- Tabulka predikčních hodnot s 95 % intervalem spolehlivosti	62
Tabulka 6	- Kontingenční tabulka počtu úspěšných, neúspěšných a neuskutečněných obchodů vzhledem k referenčnímu období.....	80

Tabulka 7 - Kontingenční tabulka vlivu odhadnutých parametrů ARMA na úspěšnost modelu.	81
Tabulka 8 - Kontingenční tabulka vlivu p-hodnoty Ljung-Box testu rozdělených do deseti kategorií na úspěšnost obchodů	82
Tabulka 9 - Kontingenční tabulka vlivu p-hodnoty Ljung-Box testu rozdělených do dvou kategorií na úspěšnost obchodů	82

Seznam grafů

Graf 1 - Produkce platiny podle států v metrických tunách za rok 2017.....	35
Graf 2 - Spotřeba platiny dle odvětví za rok 2017 v tisících oz.	36
Graf 3 - Poptávka po zlatu v roce 2017 podle odvětví	38
Graf 4 - Zavírací denní ceny platiny v USD	56
Graf 5 - První diference denních cen platiny v USD	57
Graf 6 - ACF reziduí odhadnutého modelu	60
Graf 7 - PACF reziduí odhadnutého modelu	60
Graf 8 - Grafické znázornění predikce diferencovaných hodnot.....	61
Graf 9 - Znázornění predikce na původních datech.....	63
Graf 10 - Porovnání B&H strategie s ARIMA/GARCH strategií na zlatu.....	65
Graf 11 - Porovnání B&H strategie s ARIMA/GARCH strategií na platině.....	65
Graf 12 - Porovnání B&H strategie s ARIMA/GARCH strategií na ropě	66
Graf 13 - Hodnota 1 bodu na 1 lot pozice.....	68
Graf 14 - Minimální bodové pohyby na vybraných komoditách.....	69
Graf 15 - Velikosti spreadu u vybraných komodit.....	69
Graf 16 - Vývoj kapitálu při testování strategie na reálných datech.....	72
Graf 17 - Bodový výsledek obchodování zlata na reálných datech.....	74
Graf 18 - Bodový výsledek obchodování zlata na reálných datech s potlačením spreadu .	74
Graf 19 - Bodový výsledek obchodování ropy na reálných datech	76
Graf 20 - Bodový výsledek obchodování ropy na reálných datech s potlačením spreadu .	76
Graf 21 - Bodový výsledek obchodování platiny na reálných datech	78
Graf 22 - Bodový výsledek obchodování platiny na reálných datech s potlačením spreadu	78

Seznam použitých zkratek

ACF	Autokorelační funkce
AIC	Akaikeho informační kritérium
AOS	Automatický obchodní systém
API	Aplikační programovací rozhraní
AR	Autoregresivní proces
ARCH	Základní model podmíněné heteroskedasticity
ARIMA	Integrovaný smíšený proces ARMA
ARMA	Smíšený proces AR a MA
B&H	Buy and Hold investiční strategie
CFD	Rozdílový kontrakt
GARCH	Generalizovaný autoregresivní model podmíněné heteroskedasticity
HW	Hardware
MA	Proces klouzavých součtů
MT4	MetaTrader4
OTC	Typ uspořádání trhu, kde obchodování neprobíhá pod záštitou oficiální instituce
PACF	Parciální autokorelační funkce
RRR	Vztah zisku k riziku
S/R zóna	Zóna supportu či rezistence grafu – místa opakovaných odrazů
USD	Dolar

1. Úvod

Obchodování s CFD kontrakty je v dnešní době velmi populární záležitostí. K obchodování již není třeba stát na parketu burzy. Na zadávání obchodních příkazů stačí připojení k internetu a obchodní platforma. Dokonce není třeba ani vysoký kapitál, jelikož všichni obchodní zprostředkovatelé (brokeři) poskytují na obchodování finanční páku, která násobně zvyšuje dostupnost obchodního účtu. Obchodní platforma dokonce umožňuje naprogramování si vlastního nástroje pro usnadnění rozhodování, zda uzavřít, či neuzavřít obchod.

Většina obchodníků k rozhodování o otevření obchodní pozice používá vizuální analýzu grafu a standardní indikátory, z nichž některé nejvýznamnější jsou představeny v této práci. K analýze ekonomických časových řad a predikcím se velmi často používají ARIMA/GARCH modely, které mají uplatnění zejména ve finančnictví, pojišťovnictví a ekonomii. To byl námět k vytvoření takového obchodního systému, který není založen na čtení z grafu a standardních indikátorech, nýbrž na matematicko-statistických modelech, které jsou vytvářeny právě pro odhadování budoucího vývoje. Cílem práce je tedy využít možnost naprogramování si vlastního indikátoru a vytvořit obchodní systém založený na ARIMA/GARCH modelech, pomocí kterého je možné obchodovat CFD komoditní kontrakty.

Pro obchodování byl vybrán právě komoditní trh, jelikož komodity mají silné trendy, které by právě zvolené matematicko-statistické modely mohly odhalit a zužitkovat. Zároveň odhadují, kdy se trh otočí opačným směrem.

Práce je rozdělena do několika stěžejních kapitol. První a druhá kapitola se věnuje úvodu do problematiky, metodice zpracování a výpočtů ARIMA/GARCH modelů. Ve třetí kapitole jsou popsány teoretická východiska pro obchodování s komoditami na burze, která jsou potřebná pro vytvoření obchodního systému. Zejména je zde rozepsána specifikace vybraných komodit, funkce komoditních burz, možnosti analýz trhu, funkce indikátorů, dále podkapitola zabývající se risk managementem. Poslední část této kapitoly je zde věnována právě obchodním systémům. Další – čtvrtou kapitolou je vlastní práce a samotné zpracování obchodního systému. V této kapitole je nejprve provedena analýza v programu R pro vytvoření kvalitního algoritmu, který bude schopen predikovat časové řady. Následně je

vytvořen backtest v programu R, jehož úspěšnost se porovnává s investiční strategií Buy and Hold. Po úspěšném vytvoření backtestů je provedena implementace algoritmu R do platformy Meta Trader 4, na které je prováděn měsíční test na reálných datech. V páté kapitole práce je vyhodnocení výsledků obchodování na reálných datech a to, jak z celkového pohledu, tak hodnocení výsledků obchodování jednotlivých komodit. Šestá kapitola obsahuje celkové shrnutí všech zjištěných poznatků a doporučení vyplývajících z této práce.

2. Cíl práce a metodika

2.1 Cíl práce

Hlavním cílem práce je vytvořit obchodní systém, který bude obchodovat s vybranými komoditami, za využití matematicko-statistických modelů ARIMA a GARCH.

Díličními cíli jsou:

- Vytvoření kvalitních predikcí ceny vybraných komodit na základě ARIMA/GARCH modelů v R
- Verifikace predikcí formou backtestu v prostředí R
- Implementace obchodního systému do prostředí MetaTrader4
- Ověření vytvořeného systému na reálných datech.

2.2 Metodika

Těžiště diplomové práce spočívá v tvorbě obchodního systému, který bude programován v jazyce MQL4 na platformě MetaTrader4, jehož hlavní výpočty budou programovány v jazyce R. K tvorbě systému budou využity modely GARCH, ARIMA. Funkčnost systému bude dále ověřována na reálných datech po dobu jednoho měsíce v prostředí MetaTrader4. O konkrétní metodice výpočtů ARIMA/GARCH modelů pojednává následující kapitola.

2.2.1 Metodika výpočtů ARIMA/GARCH modelů

2.2.1.1 Úvod do časových řad

Co se týče lineárních i nelineárních modelů, ze všeho nejdříve si je třeba ujasnit pár klíčových pojmů. Zejména stochastický proces, definici časové řady a stacionaritu dat, jelikož ta je nezbytná pro výpočty lineárních i nelineárních modelů.

Stochastický proces (náhodný proces) je množina náhodných veličin $\{X(s, t), s \in S, t \in T\}$, kde S je výběrový prostor a T je indexová řada. Pro každé $t \in T$ je $X(., T)$ náhodná veličina definovaná na výběrovém prostoru S a pro každé $s \in S$ je $X(s, .)$ Realizace

stochastického procesu definována na indexové řadě T . To znamená, že $X(s, t)$ je náhodná veličina závislá na čase [1][2]. Stochastické procesy jsou zatíženy neurčitostí, na rozdíl od deterministických procesů, které se dají jednoznačně popsat matematickou rovnicí. Stochastický přístup umožňuje popsat výskyt extrémů v časových řadách, což je velmi důležité v praxi téměř ve všech odvětvích. V současnosti nachází nejčastěji uplatnění v ekonomické oblasti [1][3] a to ve finančnictví a pojišťovnictví.

Časová řada je chronologicky uspořádána množina dat. Protože tato práce pojednává o analýze časových řad se stochastický přístupem, budeme v následujícím textu časovou řadu $\{y_t, t = 1, 2, \dots, n\}$ chápat jako konkrétní realizaci stochastického procesu $\{Y_t\}$. Y_t je náhodná proměnná s vlastním rozdělením pravděpodobnosti a každé y_t je jedna hodnota této proměnné. [4]

Stacionarita dat je nezbytná pro to, aby bylo možné v praxi analyzovat náhodný proces. Analýza nestacionárního procesu je přesto možná. A to v případě, pokud bychom nestacionární proces diferencovali na dílčí procesy, které už jsou stacionární nebo kvazistacionární [2]. Tuto možnost zatím nebudeme brát v úvahu a budeme uvažovat pouze stacionární náhodný proces.

Striktní stacionarita (nebo stacionarita v užším smyslu), předpokládá, že rozdělení náhodného procesu je invariantní vůči posunu v čase, tj jeho úplný statistický popis se nemění. Protože tuto podmínku není jednoduché v praxi dodržet, byl zaveden pojem slabá (kovarianční) stacionarita.

Slabá stacionarita (také stacionarita v širším smyslu) je naopak méně omezující a požaduje, aby měl příslušný náhodný proces konstantní momenty druhého řádu, tj střední hodnotu a rozptyl. Kovarianční a korelační funkce slabě stacionárního procesu závisí pouze na časové vzdálenosti (posunutí) náhodných veličin, a ne na umístění v časové řadě [1][3]. Časovou vzdálenost k ve spojení s autokorelační funkcí nazýváme řádem a má významnou roli při určení správného modelu. Mimo jiné je omezení se na kovariance stacionárního (časově invariantního) náhodného procesu výhodné i z důvodu zjednodušení vztahu teoretických charakteristik. To znamená, že pro $t = 0, 1, \dots, n$ platí, že:

$$\text{a) střední hodnota} \quad \mu_t = E(y_t) = \mu \quad (1)$$

$$\text{b) rozptyl} \quad \sigma_t^2 = \text{var}(y_t) = D(y_t) = E[(y_t - \mu_t)^2] = \sigma^2 \quad (2)$$

c) autokovariční funkce řádu k ($k = 0, \pm 1, \dots$)

$$\gamma_k = \gamma(t, t - k) = \text{cov}(y_t, y_{t-k}) = E[(y_t - \mu)(y_{t-k} - \mu)] \quad (3)$$

d) autokorelační funkce řádu k ($k = 0, \pm 1, \dots$)

$$\rho_k = \frac{\text{cov}(y_t, y_{t-k})}{\sigma^2} = \frac{\gamma_k}{\gamma_0} \quad (4)$$

2.2.1.2 Autokorelační funkce – autokorelace reziduí

Reziduální složku s označením a_t , $t = 1, 2, \dots, n$ představuje **bílý šum** s normálním rozdělením $N(0, \sigma_a^2)$. To znamená, že jde o stacionární náhodný proces tzv. IID proces (proces s nezávislým rozdělením pravděpodobnosti), v němž a_t , $t = 1, 2, \dots, n$ jsou nekorelované proměnné s obvykle nulovou střední hodnotou a s konstantním rozptylem σ_a^2 . Autokovarianční funkce takového procesu je také rovná nule. Zmíněné vlastnosti jsou matematicky vyjádřené podmínkami: (14), (15), (16).

Speciálním případem je gaussovský šum, proces s normálním rozdělením $N(0, \sigma_a^2 = 1)$. V modelech reziduální složka reprezentuje například chyby v měření, špatnou volbu regresního modelu a různé jiné vlivy. V odborné literatuře se pro tuto složku používá různé označení, nejčastěji však epsilon, v této práci budeme z důvodu přehlednosti využívat označení a_t . [4]

Reziduální autokorelační funkce

Obdobně, jako autokorelační funkce i reziduální autokorelační funkce vyjadřuje míru lineární závislosti časově opožděných veličin a_t a a_{t-k} . Autokorelace reziduí je definována vztahem

$$r_k = \hat{\rho}_k = \frac{\sum_{t=k+1}^T \hat{a}_t \hat{a}_{t-k}}{\sum_{t=1}^T \hat{a}_t^2} \quad (5)$$

Pomocí reziduálních (nesystematických) složek a_t můžeme modelovat jednoduchý typ autokorelace

$$y_t = a_t + \rho_1 a_{t-1} \quad (6)$$

Parametr ρ_1 nabývá hodnoty z intervalu $-1, 1$. Hodnoty blízké 1 napovídají, že jde o pozitivní korelaci. Naopak hodnoty blízké -1 signifikuje zápornou korelaci. V případě, že

rezidua nejsou autokorelovaná, nabývá reziduální autokorelační funkce, respektive parametr ρ_1 , hodnotu rovnou 0. Vztah (2.2) popisuje proces klouzavých součtů 1. řádu (MA(1), proto ρ_1). Autokorelace reziduí je v případě tvorby matematicko-statistických modelů nežádoucí. [1]

2.2.1.3 Lineární proces a lineární modely

Lineární proces můžeme definovat jako nekonečnou řadu:

$$y_t = a_t + \Psi_1 a_{t-1} + \Psi_2 a_{t-2} + \dots, \quad (7)$$

kde ψ_i jsou vzájemně nekorelované parametry a složky a_t reprezentují bílý šum s nulovou střední hodnotou a rozptylem σ_a^2 . Při zápisu lineárního procesu se často používá operátor zpětného posunutí B, pro který platí $B^j y_t = y_{t-j}$. Totéž platí pro složku a_t . S pomocí operátoru B se dá zapsat vztah (3.1) ve tvaru $y_t = \psi(B)a_t$, kde

$$\Psi(B) = 1 + \Psi_1 B + \Psi_2 B^2 + \dots = 1 + \sum_{j=1}^{\infty} \Psi_j B^j \quad (8)$$

Podmínka stacionarity: Lineární proces existuje, pokud platí, že $\psi(B)$ konverguje pro $|B| \leq 1$, B je obyčejná číselná proměnná. Tato podmínka zaručuje stacionaritu lineárního a nulovost jeho střední hodnoty $E(y_t) = 0$.

Praxe vyžaduje, aby se současná hodnota y_t lineárního procesu dala vyjádřit pomocí hodnot předchozích y_{t-1}, y_{t-2}, \dots a současné hodnoty bílého šumu a_t ve tvaru

$$y_t = \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + \dots + a_t \quad (9)$$

a pomocí operátoru posunutí B můžeme (9) zapsat $\Phi(B)y_t = a_t$, kde

$$\Phi(B) = 1 - \Phi_1 B - \Phi_2 B^2 - \dots = 1 - \sum_{j=1}^{\infty} \Phi_j B^j \quad (10)$$

Lineární procesy, které se dají takto vyjádřit se nazývají **invertibilní**. **Podmínka invertibility:** Lineární proces je invertibilní, jestliže $\Phi(B)$ konverguje pro $|B| \leq 1$. Platí mezi parametry ψ_j a Φ_j získáme z podmínky $\psi(B)\Phi(B) = I$. Při výpočet pracujeme s $\psi(B)$ a $\Phi(B)$ jako s mocninnými řadami. Po dosazení (9) a (10) dostaneme

$$1 + (\Psi_1 - \Phi_1)B + (\Psi_2 - \Phi_1\Psi_1 - \Phi_2)B^2 + \dots = 1, \quad (11)$$

po úpravě (11) získáme $\psi_1 = \Phi_1$, $\psi_2 - \Phi_1\psi_1 - \Phi_2 = 0$, ... analogicky.

V Box-Jenkinsové metodologii mají praktický význam procesy s konečným, co nejmenším počtem nenulových parametrů. **Parametry se volí tak, aby vyhovovaly podmínkám pro stacionaritu a invertibilitu.** [1]

Jako prvním lineárním modelem, který je třeba si objasnit je **Proces klouzavých součtů MA(q)**. Proces klouzavých součtů je často označován jako proces klouzavých průměrů (Moving Average), v odborné terminologii se používá zkratka MA(q).

$$y_t = a_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q} = \theta_q(B)a_t \quad (12)$$

Vztah (3.6) popisuje proces klouzavých součtů řádu q .

$$\text{Mocninná řada} \quad \theta_q(B) = 1 + \sum_{j=1}^q \theta_j B^j \quad (13)$$

představuje operátor klouzavých součtů, při analýze odhadujeme parametry procesu $\theta_1, \theta_2, \dots, \theta_q$, které jsou reálná čísla.

Pro složky bílého šumu a_t platí:

$$E(a_t) = 0, \text{ pro } \forall t, \quad (14)$$

$$D(a_t) = E(a_t^2) = \sigma_a^2 > 0, \quad (15)$$

$$\text{cov}(a_t, a_s) = E(a_t a_s) = 0, \text{ pro } t \neq s \quad (16)$$

Samotný proces MA(q) má tyto vlastnosti [1]:

- 1) Je vždy stacionární, $\theta_q(B)$ konverguje pro libovolné hodnoty parametrů.
- 2) Střední hodnota je nulová $E(y_t) = 0$
- 3) Pro rozptyl platí: $\text{var}(y_t) = \sigma_a^2(1 + \sum_{j=1}^q \theta_j^2)$ (17)
- 4) ACF má tvar $\rho_k = \frac{\theta_k + \theta_1\theta_{k+1} + \dots + \theta_{q-k}\theta_q}{(1 + \sum_{j=1}^q \theta_j^2)}$. Identifikační bod je $k_0 = q$. (18)
- 5) Parciální autokorelační funkce f_{kk} procesu MA(q) nemá identifikační bod.

- 6) Proces $MA(q)$ je invertibilní, pokud všechny jeho nulové body, kořeny polynomu $\theta_q(B) = 0$, leží uvnitř jednotkové kružnice [1]. To znamená, že pro kořeny tohoto polynomu musí platit $|\theta_q| < 1$.

Z uvedených vlastností je důležité zapamatovat si 1) a 6). Z těchto vlastností vyplývá, že stacionarita $MA(q)$ je zaručena. Naopak invertibilitu musíme při stanovení modelu otestovat.

Dalším ze základních lineárních procesů je **autoregresivní proces $AR(p)$** . Autoregresivní proces řádu p označovaný jako $AR(p)$ je definován tvarem

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + a_t \quad (19)$$

Proces $AR(1)$ má jen první parametr ϕ_1 . Ten můžeme chápat jako indikátor paměti procesu, čímž je jeho absolutní hodnota bližší hodnotě 1, tím bude paměť delší a naopak, čím bude ϕ_1 bližší nule, tím bude paměť kratší. Pokud je roven nule, ACF je nulová, proces nemá paměť a tím pádem jde pouze o samotný proces šumu a_t .

S použitím operátoru zpětného posunutí můžeme napsat

$$\phi(B)y_t = a_t, \quad (20)$$

kde
$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \quad (21)$$

je autoregresivní operátor.

Základní vlastnosti procesu $AR(p)$ [1][3]:

- 1) $AR(p)$ je stacionární, pokud všechny kořeny polynomu $\phi_p(B)$ leží uvnitř jednotkové kružnice.

- 2) Pro střední hodnotu platí $E(y_t) = 0$

3) Pro rozptyl platí $var(y_t) = \frac{\sigma_a^2}{1 - \phi_1 \rho_1 - \phi_2 \rho_2 - \dots - \phi_p \rho_p} \quad (22)$

- 4) ACF procesu $AR(p)$ splňuje soustavu diferenčních rovnic

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p}, \text{ pro } k > 0. \quad (23)$$

- 5) PACF f_{kk} má identifikační bod $k_0 = p$, z čehož vyplývá, že $f_{kk} = 0$ pro $k > p$.

- 6) Proces $AR(p)$ je invertibilní pro libovolné hodnoty parametrů.

Se znalostí procesu klouzavých součtů $MA(q)$ a autoregresivních procesů $AR(p)$, můžeme oba procesy propojit a vznikne tak **smíšený proces ARMA (p,q)** . Tyto modely se v praxi používají kvůli různým vlastnostem jednotlivých procesů / modelů.

Proces řádu p a q je definován vztahem

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q} \quad (24)$$

nebo pomocí operátoru zpětného posunutí.

$$\phi_p(B)y_t = \theta_q(B)a_t \quad (25)$$

kde $\Phi(B)$ je autoregresivní operátor a $\theta(B)$ operátor klouzavých součtů. Vlastnosti smíšeného procesu se dají odvodit z vlastností dílčích procesů $AR(p)$ a $MA(q)$.

- 1) $ARMA(p, q)$ je stacionární, pokud je stacionární proces $AR(p)$. Tedy platí $|\phi_1| < 1$
- 2) Střední hodnota stacionárního modelu $ARMA(p, q)$ je nulová.
- 3) ACF i PACF vyhovuje soustavě diferenčních rovnic (23) pro $k > q$. Nemá identifikační bod a po prvních q a p hodnotách představuje lineární kombinaci klesajících geometrických posloupností a sinusoid různých frekvencí s klesajícími amplitudami.
- 4) $ARMA$ proces je invertibilní, pokud je invertibilní proces $MA(q)$ $|\theta_1| < 1$.

V praxi se nejčastěji setkáme s modelem $ARMA(1,1)$. Z definičního vztahu (24) je zřejmé, že pokud při procesu $ARMA(1,1)$ $\phi_1 = 0$, proces se redukuje na proces $MA(1)$ a naopak, pokud $\theta_1 = 0$, $ARMA(1,1)$ se redukuje na proces $AR(1)$. Proces $ARMA(1,1)$ je prostá pokud $|\theta_1| < 1$ a stacionární pokud $|\phi_1| < 1$. [1]

Pakliže řada není stacionární, není možné aplikovat modely $ARMA$. Právě proto jsou tu **integrované procesy $ARIMA(p,d,q)$** , které umožňují modelovat nestacionární časové řady. Je však nutné, aby se daly převést na stacionární. To se realizuje diferencováním na dílčí řady, které již jsou stacionární. Smíšený integrovaný model $ARIMA(p, d, q)$ se popisuje vztahem

$$\phi_p(B)\Delta^d y_t = \theta_q(B)a_t \quad (26)$$

$\Delta^d y_t$ představuje časovou řadu zkonstruovanou diferenciací původní časové řady a d vyjadřuje řád této diferenciaci. $\Delta = (1-B)$ je diferenční operátor. Můžeme zapsat:

$$\Delta^d y_t = y_t(1 - B)^d \quad (27)$$

Pokud bychom použili diferenciaci například řádu $d = 1, 2, \dots$ dostali bychom pro $d = 1$ $\Delta y_t = y_t - y_{t-1}$, a pro $d = 2$, $\Delta^2 y_t = y_t - 2y_{t-1} + y_{t-2}$ a analogicky.

Sestavení takového modelu se realizuje ve dvou krocích:

- 1) Nejdříve se původní časová řada y_t zdiferencuje na stacionární řadu $\Delta^d y_t$. Diferencováním se původní časová řada zkrátí.
- 2) Jako další krok se aplikuje smíšený model ARMA(p, q) [1]

Pro účely této práce byla na všechny časové řady aplikovaná diferenciaci prvního řádu, jež je možné aplikovat na většinu pozorování. V programu R se ARIMA spočítá velice jednoduše pomocí funkce `arima()`, které se předá časová řada a parametry, pro které se má ARIMA spočítat. Výsledný kód pro spočtení funkce ARIMA bude vypadat následovně: `arima(returns, order = c(p, 0, q))`, kde `returns` je diferencovaná časová řada, pro kterou se má model spočítat a vektor `c` obsahuje parametry modelu.

2.2.1.4 Nelineární proces a nelineární modely

Nelineární proces lze popsat následující strukturou [5]:

$$y_t = g(a_{t-1}, a_{t-2}, \dots) + a_t h(a_{t-1}, a_{t-2}, \dots) \quad (28)$$

kde **střední podmíněná hodnota** tohoto procesu je

$$E[y_t | y_{t-1}] = g(a_{t-1}, a_{t-2}, \dots) \quad (29)$$

a **podmíněný rozptyl**

$$var[y_t | y_{t-1}] = E[\{(y_t - E[y_t]) | y_{t-1}\}^2] = \{h(a_{t-1}, a_{t-2}, \dots)\}^2 \quad (30)$$

Funkce $g(\cdot)$ Odpovídá podmíněné střední hodnotě procesu y_t , a funkce $h(\cdot)$ je koeficientem úměrnosti mezi změnou y_t a a_t [5]. Ze vztahu (29) vyplývá, že pokud je $g(\cdot)$

nelineární, tak potom je model nelineární ve střední hodnotě. A podle vztahu (30) pokud je nelineární funkce $h(\cdot)$, pak je i model nelineární v rozptylu.

Při určení podmíněného rozptylu vycházíme z teorie pravděpodobnosti, konkrétně podmíněné pravděpodobnosti [6], kde náhodné proměnné jsou na sobě závislé. Mějme dvě závislé náhodné proměnné X a Y . Potom pro rozptyl proměnné X podmíněné Y platí:

$$\text{var}(X|Y) = E[X - E[(X|Y)]|Y] \quad (31)$$

V případě, že by X a Y byly nezávislé:

$$\text{var}(X|Y) = \text{var}(X). \quad (32)$$

Klasický lineární regresní model je charakterizován požadavkem konečného a konstantního rozptylu náhodných složek $D(y_t) = \sigma^2$ pro $\forall t$. Tento požadavek platí také pro rezidua modelu, $D(a_t) = \sigma^2$ pro $\forall t$, odborně nazývá homoskedasticita. Proces s konečným a konstantním rozptylem složek můžeme označit jako **homoskedastický**. Naopak v případě náhodného výběru pozorování, ve kterém dochází k velkým změnám hodnot (vysoká variabilita) mluvíme o opačném jevu o **heteroskedasticitě** [7]. V takovém případě je porušena podmínka a můžeme zapsat $D(a_t) \neq \sigma^2$ pro $\forall t$.

Za příčiny výskytu heteroskedasticity můžeme označit samotné pozorování (v našem případě sledovaná časová řada). Významnou roli hraje i existence chyb měření.

Jinou častou příčinou je nesprávná specifikace modelu, spočívající ve vynechání významné proměnné, nebo v použití skupinových průměrů místo původních pozorování [7]. Ve většinu případů je však těžké určit příčinu heteroskedasticity, proto se ji snažíme odstranit různými transformacemi. V případě výskytu heteroskedasticity v samotné podstatě dat, je vhodnější při procesu vytváření modelu zohlednit tuto vlastnost a využít modely, které jsou schopné zachytit heteroskedasticitu. Heteroskedasticita je pojem pro proměnný a nekonečný rozptyl náhodných poruch a reziduí v čase.

Jako první popsany nelineární model je **proces ARCH(q)**. Poslední léta vývoje různých odvětví, ve kterých se aplikuje analýza časových řad, podněcují k využívání nelineárních modelů. Předpokládá se, že i zdánlivá nezávislost volatilních jevů (nestálých nebo s velkou variabilitou) může být pomocí analýzy časových řad predikovaná na základě

existující nelineární závislosti. Ukázalo se, že ARCH modely jsou vhodným nástrojem pro takovou analýzu. Zkratka ARCH pochází z anglického názvu pro autoregresivní podmíněnou heteroskedasticitu (Autoregressive Conditional Heteroskedasticity). Zavedl je v roce 1982 Engle [5].

Pokud si všimneme například vztahu (19) pro lineární proces $AR(p)$, zbytková složka a_t je aditivně přidána k inovacím. Protože a_t má konstantní a konečný rozptyl (15), $AR(p)$ model je schopen úspěšně modelovat proces pouze pod podmínkou jeho konstantního rozptylu.

ARCH model však dovoluje modelovat změnu podmíněného rozptylu a změny minulých reziduí v čase. Kromě toho rozlišuje mezi podmíněným rozptylem a nepodmíněným rozptylem. Nepodmíněný rozptyl se na rozdíl od podmíněného zachovává konstantní. Tyto vlastnosti umožňují částečně popsat nelinearitu, poradit si s hromaděním chyb, a především predikovat náhlé změny. [5]

Podle Engleho se ARCH proces matematicky definuje vztahem:

$$y_t = \sqrt{\alpha a_{t-1}^2}, \quad (33)$$

kde funkce $g(\cdot) = 0$, a $h(\cdot) = \sqrt{\alpha a_{t-1}^2}$, proto je tento model nelineární v rozptyle, ale je lineární ve střední hodnotě [5].

Na základě výše definovaného procesu (33) navrhl Engle následující model, který má za úkol zachytit sériovou korelaci jevů s vysokou variabilitou: [5]

$$\sigma^2 = \omega + \alpha(L)\eta^2, \quad (34)$$

kde $\alpha(L)$ označuje polynom reprezentující opožděné parametry, $\eta_t | \Psi_{t-1}$ jsou inovace s rozdělením $\sim N(0, \sigma_{t-1}^2)$. ARCH model charakterizuje distribuci stochastických odchylek a_t podmíněných skutečnými hodnotami časové řady proměnných $\Psi_{t-1} = \{y_{t-1}, x_{t-1}, y_{t-2}, x_{t-2}, \dots\}$.

Po úpravách v (34) můžeme zapsat ekvivalentní reprezentaci $ARCH(q)$ modelu takto:

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i a_{t-i}^2 \quad (35)$$

V tomto případě pro nesystematické složky a_t pro $\forall t$ platí

$$a_t = \eta_t \sqrt{\sigma_t^2}, \quad (36)$$

tj. mají proměnlivý (nekonstantní) rozptyl. Označení η_t je pro inovace, které mají charakter šumu $\eta_t = N(0, \sigma_t^2) \sim N(0,1)$, kde rozložení $N(0,1)$ indikuje Gaussovský šum. Ze vztahu (35) je vidět časovou závislost rozptylu.

Podle literatury [3][5], je možné bez ztráty obecné platnosti definovat jednoduchý ARCH model:

$$y_t = a_t a_{t-1}^2, \quad (37)$$

kde a_t je bílý šum.

Co se týče vlastností procesu (37), popíšeme si je v následujících výpočtech: [3]

$$E(y_t) = E(a_{t-1}^2 a_t) = E(a_{t-1}^2) E(a_t) = 0 \quad (38)$$

$$\text{var}(y_t) = E(y_t^2) = E(a_{t-1}^4 a_t^2) = E(a_{t-1}^4) E(a_t^2) = (3\sigma_a^4) \sigma_a^2 = 3\sigma_a^6 \quad (39)$$

$$E(y_t | y_{t-h}) = E(a_{t-1}^2 a_t a_{t-h-1}^2 a_{t-1}) = E(a_t) E(a_{t-1}^2 a_{t-h-1}^2 a_{t-1}) = 0, \text{ pro } h > 0 \quad (40)$$

* Využili jsme poznatek o normovaném momentu 4. řádu $E(a_t^2)$ nazývaného koeficient špičatosti [1], u kterého při normálním rozdělení a_t platí

$$m_4^*(a_t) = E\left(\frac{a_t^4}{\sigma_a^4}\right) = \frac{E(a_t^4)}{\sigma_a^4} = 3 \rightarrow E(a_t^4) = 3\sigma_a^4 \quad (41)$$

Z uvedených vztahů (38), (39), (40) jsou zřejmé vlastnosti procesu y_t :

- kovariančně (slabě) stacionární
- y_t nemá podmíněný rozptyl $\text{var}(y_t) = 3\sigma_a^6$ (nezávisí na čase)
- y_t má podmíněný rozptyl hodnotou y_{t-1}

$$\text{var}(y_t | y_{t-1}) = \text{var}(a_{t-1}^2 a_t | y_{t-1}) = \sigma_a^2 a_{t-1}^4 \quad (42)$$

na čase závisí prostřednictvím faktoru a_{t-1}^4 . Vztah (42) dokazuje tvrzení o nelinearitě ARCH modelu v rozptylu, které bylo zmíněno při definici.

Druhým nelineárním modelem, který je důležitý je **proces GARCH(p, q)**. Velkou nevýhodou ARCH(q) modelu je totiž to, že rostoucím řádem p vzrůstá náročnost výpočtu, proto se v praxi častěji používá generalizovaný autoregresivní podmíněný heteroskedastický model GARCH(p, q) [3]:

$$\sigma_t^2 = \omega + \alpha(L)\eta_t^2 + \beta(L)\sigma_{t-1}^2 \quad (43)$$

Pokud zvážíme podobnost struktury s modelem ARMA(p, q), tak v modelu GARCH(p, q) polynom $\alpha(L)$ řádu " q " odpovídá klouzavému součtu MA(q) a polynom $\beta(L)$ řádu " p " koresponduje s prvkem autoregresivního procesu AR(p). Lepší pochopení vztahu (43) poskytuje obdobný zápis (44).

GARCH(p, q) proces si můžeme vysvětlit následovně: Necht' a_t označuje v čase diskrétní náhodný proces s reálnými hodnotami a Ψ_t označuje informační soubor (sérii odchylek oboru σ) v čase t . GARCH(p, q) proces je pak dán následujícími vztahy: [1]

$$a_t | \Psi_{t-1} \sim N(0, \sigma_t^2),$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i a_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 = \alpha_0 + \alpha(L)a_t^2 + \beta(L)\sigma_t^2, \quad (44)$$

kde: $p \geq 0, q > 0$

$$\alpha_0 > 0, \alpha_i \geq 0, \text{ kde } i = 1, \dots, q$$

$$\beta_i \geq 0, \text{ kde } i = 1, \dots, p$$

Jestliže $p = 0$, proces se redukuje na ARCH(q) proces. V případě, že bude $p = q = 0$, zůstane nám pouze jednoduchý bílý šum (složka a_t). Zatím, co v ARCH(q) procesu je podmíněný rozptyl specifikovaný pouze jako lineární funkce minulých vzorků podmíněného rozptylu, GARCH (p, q) proces zavádí i vliv opožděných podmíněných rozptylů. Tato vlastnost se dá vnímat jako druh adaptivní učícího se mechanismu.

Proces definovaný v (44) je stacionární s vlastnostmi: [8]

$$E(a_t) = 0, \text{ var}(a_t) = \alpha_0(1 - A(1) - B(1))^{-1} \text{ a } \text{cov}(a_t, a_s) = 0 \text{ pro } t \neq s \quad (45)$$

v jediném případě pokud $A(1) + B(1) < 1$.

Nejjednodušší a zároveň v praxi nejvyžívanější GARCH model je bezpodmínečně GARCH(1,1) [1][8]. Matematický zápis tohoto modelu dostaneme odvozením od obecného vzorce (44):

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \text{ kde platí } \alpha_0 > 0, \alpha_1 \geq 0, \beta_1 \geq 0. \quad (46)$$

2.2.1.5 Smíšený model ARIMA/GARCH

Na základě předchozí studií konkrétních modelů spolu se zvážением referencí v literatuře [1][8] jsem zvolil smíšený ARIMA/GARCH model jako nejvhodnější k otestování možností jeho využití při analýze a predikci. V této kapitole bude blíže popsán samotný model ARIMA/GARCH, a postup při konstrukci modelu. Následující text je souhrnem všech, doposud probraných faktů, které se týkají modelování časových řad.

Už ze samotného názvu vyplývá, že se jedná o kombinaci dvou modelů lineárního ARIMA (p, d, q) a nelineárního modelu GARCH (p, q). Model můžeme označit i jako model podmíněné střední hodnoty a podmíněného rozptylu. V následujících řádcích si tento model vyjádříme matematicky. Aby nedošlo k záměně parametrů označujících řády jednotlivých modelů, u modelu ARIMA budeme používat místo parametru p parametr r , označující řád autoregresivního procesu (AR). Označení řádu q procesu klouzavých součtů (MA) nahradíme proměnnou m . Označení řádu diferenciace d a řádů procesu GARCH p a q necháme bez změny. Zápis ARIMA modelu bude následující ARIMA (r, d, m).

Pokud t_y jsou časové řady, které chceme modelovat, potom obecný model **podmíněné střední hodnoty** (ARIMA (r, d, m)) vyjadřuje následující zápis:

$$(\nabla^d y_t) = \sum_{i=1}^r \phi_i (\nabla^d y_{t-1}) + a_t + \sum_{j=1}^m \theta_j (\nabla^d y_{t-j}), \quad (47)$$

Připomeňme si, že místo zápisu (47) můžeme použít i tvar:

$$\phi_p(B) \Delta^d y_t = \theta_q(B) a_t, \quad (48)$$

kde $a_t \sim N(0, \sigma_t^2)$, $\phi(\cdot)$ a $\theta(\cdot)$ představují polynomy r -tého a m -tého stupně, B je operátor zpětného posunutí ($B^j y_t = y_{t-j}$, $B^j a_t = a_{t-j}$, $j = 0, \pm 1, \pm 2, \dots$), hodnota d hovoří o stupni diferenciaci. proměnná t_a označuje inovace původní časové řady.

$$\nabla y_t = y_t - y_{t-1} = (1 - B)y_t \quad (49)$$

Vztah (49) definuje diferenční operátor se zpožděním "lag" = 1.

Nyní definujme podmíněný rozptyl inovací t_a , pro který platí:

$$\sigma_t^2 = v_{t-1}(y_t) = E_{t-1}(a_t^2) \quad (50)$$

V jednání o ARCH(p) procesu jsme uvedli, že podstata jeho použití spočívá ve schopnosti rozlišit podmíněný a nepodmíněný rozptyl v procese inovací t_a . Termín "podmíněný" jednoznačně naznačuje závislost na předchozích pozorováních. Zatímco termín "nepodmíněný" se týká dlouhodobého chování časové řady. Takové chování nevyžaduje zvláštní poznatky o minulých hodnotách řady. GARCH model nám tedy charakterizuje podmíněnou distribuci inovací zachycením sériové závislosti. [1]

Všeobecný GARCH(p, q) model pro podmíněný rozptyl inovací t_a je,

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \beta_i \sigma_{t-i}^2 + \sum_{j=1}^q \alpha_j a_{t-j}^2, \quad (51)$$

S omezeními:

$$\sum_{i=1}^p \beta_i + \sum_{j=1}^q \alpha_j < 1, \alpha_0 > 0, \alpha_j \geq 0, \beta_i \geq 0, j = 1, 2, \dots, q, i = 1, 2, \dots, p \quad (52)$$

Kombinací vztahů (47) a (51) dostaneme matematické vyjádření celého ARIMA(r, d, m)/GARCH(p, q) modelu:

$$(\nabla^d y_t) = \sum_{i=1}^r \phi_i (\nabla^d y_{t-i}) + a_t + \sum_{j=1}^m \theta_j (\nabla^d y_{t-j}), a_t \sim N(0, \sigma_t^2)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \beta_i \sigma_{t-i}^2 + \sum_{j=1}^q \alpha_j a_{t-j}^2 \quad (53)$$

Odhad parametrů modelu je prvním krokem jeho přizpůsobení analyzovaným datům. Celý postup odhadu, celkově stanovení pěti parametrů r, d, m, p, q vychází z Box-Jenkinsové metodologie.

Jako první krok se stanoví parametr d , jenž vyžaduje podmínka stacionarity, která předurčuje konstrukci uspokojivého modelu. Pokud je analyzovaná řada nestacionární musíme ji diferencovat. Parametr d se odhaduje na základě výpočtu ACF a PACF.

ARIMA model popisuje podmíněnou střední hodnotu, která se mění v čase, avšak při konstantním rozptylu pozorovaných dat. Řád parametrů r a m určíme také pomocí vlastností ACF a PACF. Parametr m procesu MA můžeme přímo určí z ACF, která popisuje korelaci mezi aktuálními a minulými hodnotami časové řady. PACF, která charakterizuje korelaci mezi aktuálními hodnotami a jejich změnou oproti minulosti, je vhodná k určení parametru r autoregresivního procesu.

Parametry p a q , nelineárního modelu, se klasickým způsobem určují pomocí autokorelační a parciální autokorelační funkce podobně jako u ARIMA modelu. Na základě ACF a PACF však tyto parametry nejsme schopni určit přesně. Z praxe je známo, že nastavení těchto dvou parametrů na hodnotu 1 je adekvátním řešením z hlediska zachycení podmíněného rozptylu. Znamená to, že bychom si měli vystačit s nelineárním modelem GARCH (1, 1).

Po odhadnutí parametrů r , d , m , p , q modelu následuje výpočet reálných hodnot parametrů φ_r , θ_m , α_q , β_p . Aby nedošlo k omylu budeme tyto parametry označovat pojmem koeficienty. Samotný proces výpočtu těchto koeficientů je oblast numerické matematiky. Používají se k tomu různé iterační metody. [7]

V programu R spočítáme smíšený model na základě předchozího odhadu ARIMA. Smíšený model se specifikuje funkcí `ugarchspec()`, která vytvoří smíšený model ARIMA/GARCH předáním parametrů ARIMA, nejčastějšího modelu GARCH(1,1) a předpokládaného rozdělení reziduí. Pro dokončení výpočtu smíšeného modelu se používá funkce `ugarchfit()`, která na specifikovaný model aplikuje data a celý model spočítá hybridní iterační metodou, která vyzkouší všechny dostupné iterační metody a pomocí vlastního algoritmu vybere tu nejlepší.

2.2.1.6 Informační kritéria

Často se stává, že jen podle dat, jejich grafického zobrazení, ACF, či PACF neumíme spolehlivě určit řád modelu. K identifikaci vhodného řádu modelu se proto používají

informační kritéria. Jsou založeny na hledání kompromisu mezi příliš vysokým řádem modelu, který snižuje rozptyl zbytkové složky, a příliš nízkým řádem modelu, čehož důsledkem je vysoká hodnota rozptylu zbytkové složky. Lepší model je ten, pro který je menší hodnota informačního kritéria. [9]

Existuje mnoho typů informačních kritérií. Pro účely této práce bylo použito Akaikeho informační kritérium.

Akaikeho informační kritérium AIC je jedno z nejpoužívanějších kritérií, které navrhl japonský statistik Hirotugu Akaike v roce 1973. Myšlenkou AIC je výběr parametrů modelu tak, aby se minimalizovala záporná hodnota logaritmicke funkce maximální věrohodnosti penalizována počtem parametrů v modelu. Snahou je tedy, aby bylo AIC co nejmenší. Je popsáno rovnicí ve tvaru

$$AIC = -2 \log(L) + 2p, \quad (54)$$

kde $\log(L)$ je hodnota logaritmu maximální věrohodnosti a p je počet parametrů v modelu. [9]

2.2.1.7 Testování nezávislosti

Při testování nezávislosti analyzovaných dat zjišťujeme, zda časová řada skládá z náhodných hodnot. **Box-Pierce test** navržený v roce 1970 [10] má jako nulovou hypotézu nezávislost analyzovaných dat. Box-Piercova testovací statistika má tvar

$$Q = n \sum_{k=1}^h r_k^2 \quad (55)$$

kde n je počet pozorování, h je maximální počet hodnot ACF, které uvažujeme a r_k^2 je mocnina k -té hodnoty výběrové autokorelační funkce reziduí. Testovací statistiku porovnááme χ^2 s rozdělením, které má za předpokladu platnosti nulové hypotézy $(h - m)$ stupňů volnosti, kde m je počet parametrů odhadovaného modelu. Pokud zjišťujeme nezávislost dat, nastaví se $m = 0$. Tento test se však často používá i při testování nezávislosti reziduí odhadnutého modelu.

Na testování nezávislosti se také velmi často využívá **Ljung-Boxův test** navržený v roce 1978 [11]. Používá se Ljung-boxova Q statistika na testování nulové hypotézy, která

zní, že v časovém řadě není autokorelace až po řád p . Tento test pracuje na stejném principu jako Box-Piercův test. Q statistika se rovněž porovnává s χ^2 rozdělením s $(h - m)$ stupni volnosti a má tvar

$$Q_{LB} = n(n + 2) \sum_{k=1}^p \frac{r_k^2}{n-k} \quad (56)$$

Tento test nezávislosti byl použit i v praktické části diplomové práce v prostředí R pomocí metody `Box.test()`.

3. Teoretická východiska burzovního obchodování

3.1 Komodity

V současnosti se na světových trzích obchoduje velké množství různých komodit. Pro lepší orientaci je důležité komodity rozčlenit do různých skupin na základě určitých kritérií. Pro účely této práce byly vybrány tři komodity a to, protože se mezi nimi nachází nízká korelace, což je důležité pro diverzifikaci portfolia.

3.1.1 Členění komodit

Odborná literatura [12][13], která se věnuje problematice komodit nejčastěji využívá následující členění na komodity:

1. **Energetické** – jsmo patří především ropa a ropné deriváty, zemní plyn, hnědé a černé uhlí a uran.
2. **Vzácné kovy** – to znamená zlato, stříbro a kovy platinové skupiny (platina, palladium, rhodium).
3. **Průmyslové kovy** – do této skupiny řadíme velké množství kovů, které se využívají hlavně v průmyslové výrobě. Patří sem železo, ocel, hliník, nikl, zinek, hořčík, titan, niob a další.
4. **Zemědělské** – které se dále člení na komodity:
 - a) rostlinného původu – pšenice, rýže, kukuřice, ječmen, žito, oves, cukr, kakao, káva, bavlna, tabák, pomerančový džus a podobně
 - b) živočišného původu – skot, prasata, drůbež, kůže
 - c) potravinářské – pšenice, kukuřice, kakao, káva, pomerančový džus
 - d) technické – bavlna, len, konopí
5. **Ostatní** – sem řadíme vodu, půdu, vzduch, povolenky, potaš, fosfor, drahokamy, a jiné

3.1.2 Ropa

Ropu jako surovinu, resp. komoditu, řadíme k neobnovitelným přírodním zdrojům energie. Ropa jako energetická surovina se nepřímou, někdy přímo podílí na většině globálně vyprodukované energie. Jako průmyslová surovina se využívá v chemickém průmyslu, kde

tvoří klíčový komponent při výrobě plastů, využívá se i při výrobě léčiv, podstatná je i při výrobě pesticidů a hnojiv, bez kterých by se zemědělský průmysl neobešel. Netřeba zapomínat na její využití v dopravě, kde se používá na výrobu pohonných hmot a v neposlední řadě se jako vedlejší produkt podílí na výrobě asfaltu. Z tohoto vyplývá, že s ropou se v každodenním životě nepřímo setkává každý z nás, a proto se někdy označuje také jako „černé zlato“. Moderní civilizace je v současnosti silně závislá na ropě, proto ji můžeme označit za nejstrategičtější komoditu vůbec, za kterou v současnosti neexistuje plnohodnotná alternativa. [14]

3.1.2.1 Nabídka ropy

Nabídku na trhu s ropou tvoří především její největší producenti. V roce 2017 byla celosvětová produkce ropy 92649 milionu barelů za den. V následující tabulce je uvedeno deset největších ropných producentů seřazených dle denní produkce ropy.

Tabulka 1 – Produkce ropy na den

Producenti ropy	Množství vyprodukované ropy v milionech barelů za den v roce 2017
USA	13057
Saudská Arábie	11951
Rusko	11257
Irán	4982
Kanada	4831
Irák	4520
Spojené Arabské Emiráty	3935
Čína	3846
Kuvajt	3025
Brazílie	2734

Zdroj: BP p.l.c. – Statistický přehled světové energie, 2019 [14]

V následující tabulce jsou seřazeny producenti ropy dle ověřených ropných rezerv.

Tabulka 2 – Množství ověřených ropných rezerv

Producenti ropy	Celkové množství ověřených rezerv v miliardách barelů ke konci roku 2017
Venezuela	303,2
Saudská Arábie	266,2
Kanada	168,9
Irán	157,2
Irák	148,8
Rusko	106,2
Kuvajt	101,5
Spojené Arabské Emiráty	97,8
USA	50,0
Libye	48,4

Zdroj: BP p.l.c. – Statistický přehled světové energie, 2019 [14]

Jak je vidět v tabulce č.1, více než 50 % celosvětové produkce ropy má na svědomí pouze deset států. Ty jsou z velké míry zodpovědní za tržní výkyvy i trendy.

Na prvním místě v produkci ropy ve světě je USA s denní produkcí za rok 2017 13057 milionů barelů. Druhým největším producentem ropy je Saúdská Arábie s denní produkcí 11951 a odhadovaným objemem rezerv 266.2 miliard barelů, za kterou následuje Rusko s produkcí 11257 milionů barelů denně. Kanada s produkcí 4831 milionu barelů denně se dostala na pátou příčku hned za Irán s produkcí 4982 milionů barelů, a to hlavně díky objevu nových nalezišť v roce 2015. Za ní následují v těsném závěsu Irák, pak Arabské Emiráty, Čína, Kuvajt a Brazílie. [14]

Co se týče ropných rezerv, tak na první příčce se objevuje Venezuela se rezervou 303,2 miliardami barelů. Předčila s velkým náskokem i zásoby Saudské Arábie které činí 266,2 miliardy barelů. Její obří zásoby jsou z valné části z důvodu působení místní politiky. Nemá sice takovou těžbu jako jiné státy, ale možnou změnou politického působení by hrála na ropném trhu velmi významnou roli. [14]

3.1.2.2 Poptávka ropy

Spotřeba ropy ve světě neustále roste. Na základě údajů od těžební společnosti BP byla denní spotřeba ropy na začátku roku 2013 něco málo přes 90 milionů barelů, na konci roku 2017 to již bylo přibližně 98 milionů barelů za den. K tomuto růstu přispěla hlavně zvýšená spotřeba v Číně, v Indii a Evropě. V následujících letech se předpokládá neustálé zvyšování spotřeby ropy. [14]

Spotřeba ropy, jak je již uvedeno výše, souvisí s průmyslem. USA jako nejmocnější industriální krajina jsou proto největším spotřebitelem ropy na světě. V roce 2017 byla průměrná denní spotřeba ropy v USA přibližně 19,9 milionů barelů za den. Následuje další silně industriální založená krajina, kterou je Čína. V roce 2017 byla denní spotřeba ropy v Číně přibližně 12,8 milionů barelů což byla historicky nejvyšší spotřeba ropy v této zemi. Indie jako třetí největší ropný spotřebitel, spotřebovala v roce 2017 přibližně 4,7 milionů barelů denně. Spotřeba ropy v Indii se neustále zvyšuje, podle agentury TASR stojí za růstem indické spotřeby ropy neustále zvyšující se životní úroveň. K dalším velkým spotřebitelům patří například Rusko, Japonsko, Saúdská Arábie či Brazílie. [14]

3.1.3 Platina

Tento drahý a zároveň průmyslový kov má bílou barvu s šedivým nádechem. Patří do skupiny tzv. platinových kovů. Rozpouští se – stejně jako zlato – pouze v lučavce královské. Proto je platina považována za velmi stálý kov. [15]

Platina (latinsky Platinum, chemická značka Pt) je chemicky velmi odolný drahý kov. Má vysokou hustotu a patří mezi jeden z nejtěžších kovů na světě. Její tvrdost je asi dvojnásobná oproti zlatu nebo stříbru. Nemá však tak dobré vodivé vlastnosti. Kromě klenotnictví se využívá také v dieselových motorech, sklářském průmyslu při výrobě optických vláken nebo v elektronice (osobní počítače). Jedná se o kov mimořádně odolný proti korozi.

Vybrané fyzikální a chemické vlastnosti platiny:

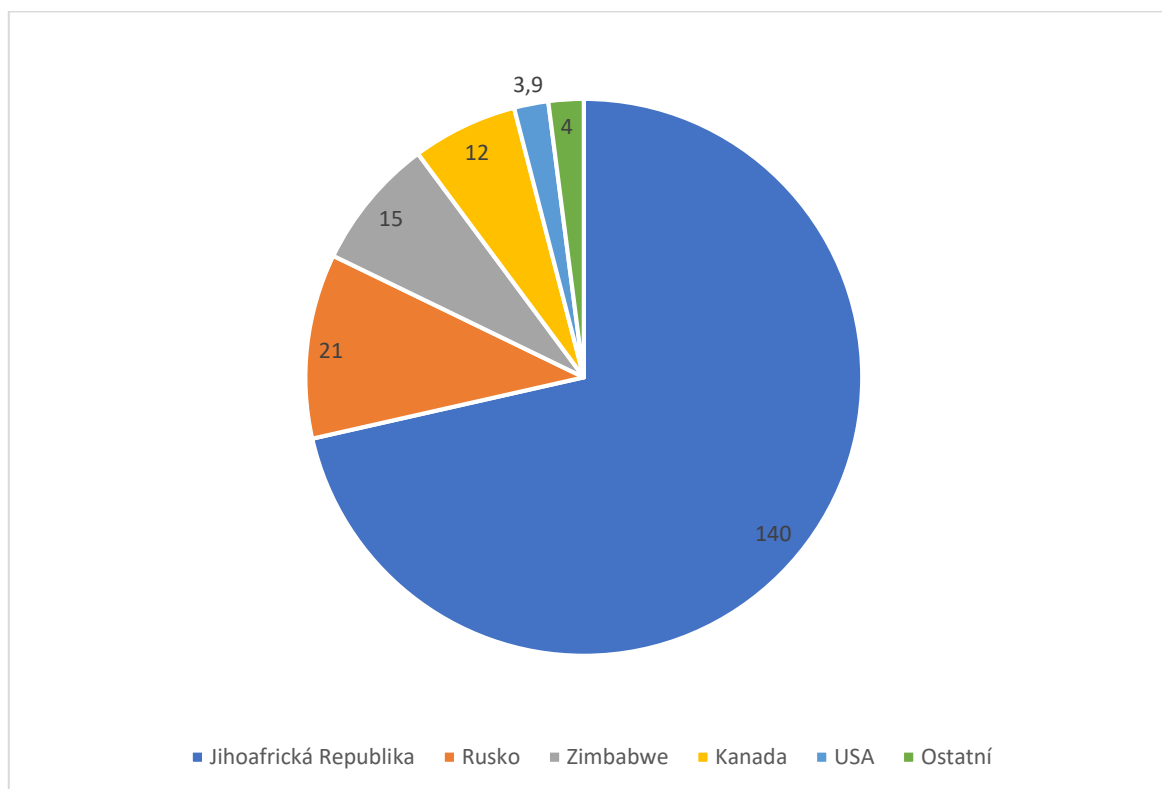
- teplota tání: 1772 °C
- hustota: 21,45 g/cm³

- elektrická vodivost: $9,4 \cdot 10^6 \text{ S m}^{-1}$
- tvrdost: 3,5
- zařazení: přechodné kovy [15]

3.1.3.1 Nabídka platiny

Těžba platiny tvoří více než tři čtvrtiny celé nabídky. Roční průměr těžby se pohybuje okolo 190 tun. Nejvíce platiny na světě se vytěží v JAR. S obrovským odstupem následuje Rusko, které produkuje množství přibližně šestkrát menší. Dále následuje Severní Amerika a Zimbabwe, které má ze zmiňovaných států největší potenciál. Množství vytěžené platiny v jednotlivých zemích je zobrazeno na níže uvedeném grafu. [16]

Graf 1 - Produkce platiny podle států v metrických tunách za rok 2017



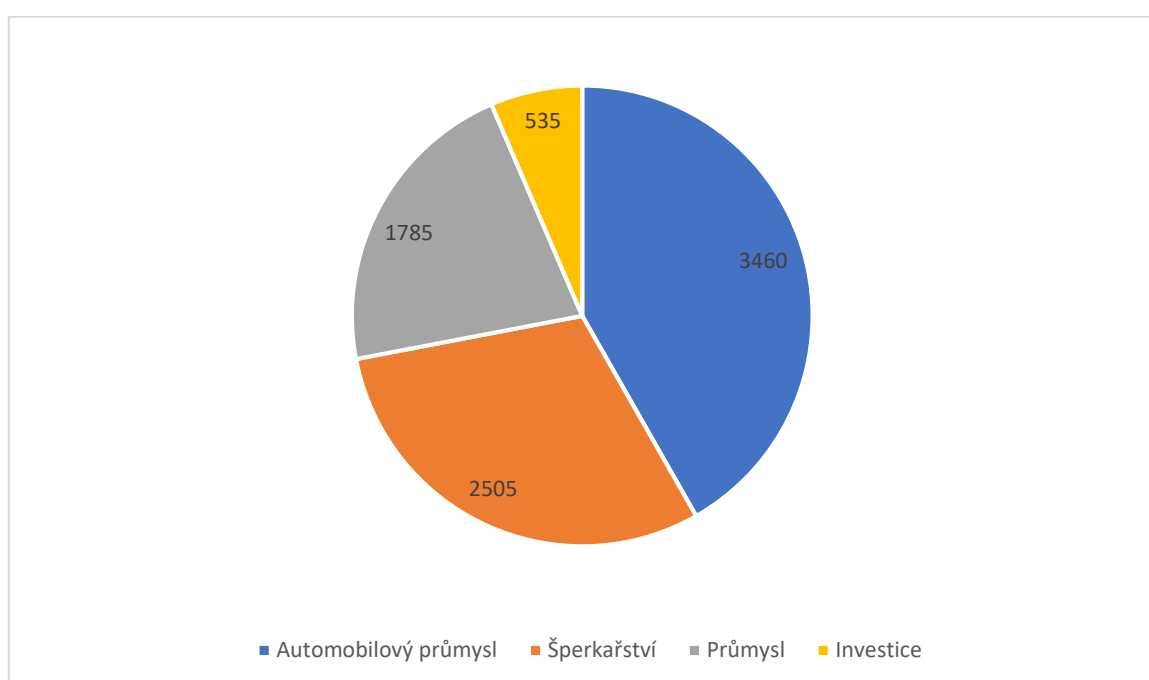
Zdroj: Statista, 2018 [16]

3.1.3.2 Poptávka platiny

Poptávka po platině je rozmanitější než její nabídka. Největší část poptávky tvoří využití platiny v katalyzátorech automobilů. Evropa spotřebuje tohoto kovu v této oblasti

poptávky nejvíce, následuje Japonsko a USA. V posledních letech vzrostla poptávka také v Číně. Tento asijský stát je velkým tahounem i v oblasti šperkařství, ve kterém zaznamenal obrovský nárůst. I v ostatních částech světa se poptávka po špercích vyrobených z platiny zvyšuje, ale oproti Číně je více než čtyřikrát menší. Téměř stejný podíl má využití tohoto drahého kovu v průmyslu. Za stěžejní obory, kde je platiny využíváno, můžeme označit chemii, elektroniku a sklářství. Poslední dva zmíněné obory jsou ale v posledních letech spíše na ústupu. [17]

Graf 2 - Spotřeba platiny dle odvětví za rok 2017 v tisících oz.



Zdroj: World Platinum Investment Council, 2018 [17]

3.1.4 Zlato

S těžbou zlata začali Egypťané už dva tisíce let před naším letopočtem. Z počátku plnilo hlavně ozdobnou funkci pro svou vizuální přitažlivost, později se stalo symbolem bohatství a uchovávatelem hodnoty. Ze zlata se razily mince a později sloužilo zlato pro krytí jmen tzv. zlatý standard. [18] Zlato jako komoditu řadíme mezi drahé kovy. Jedná se o měkký, snadno opracovatelný kov, který je vysoce odolný vůči korozi a má vysokou elektrickou a tepelnou vodivost. Díky svým vlastnostem má uplatnění zejména v elektrotechnickém průmyslu, kde se využívá na pozlacování různých spojů, či k výrobě

polovodičů. Známe je i jeho využití v lékařství např. ve stomatologii. S příměsemi jiných kovů se využívá ve šperkařství. Důležitou funkci však stále plní jako uchovatel hodnot, je součástí devizových rezerv centrálních bank. [19] Zlato je vhodné i jako investice, z dlouhodobého hlediska je to bezpečná investice hlavně v dobách recese, kdy mají centrální banky tendenci preferovat politiku tzv. „levných peněz“. Cena zlata tehdy roste. Podle Jílka patří zlato mezi nejsledovanější a nejvíce obchodované komodity na světě, neboť jako finanční nástroj je likvidní v kterékoli části světa. [20]

3.1.4.1 Nabídka zlata

Zlato v porovnání s ostatními komoditami je téměř nezničitelné, a proto všechno zlato, které kdy bylo vytěženo je stále v určité formě přítomné na zemi. World Gold Council uvádí, že podle nejlepších odhadů bylo v celé dosavadní historii vytížených přibližně 190 040 tun zlata. Z čehož přibližně 17 % tvoří rezervy v centrálních bankách, které jsou především soustředěny ve vyspělých ekonomikách západní Evropy a Severní Ameriky. [20] Nejvíce zlata v rezervách 8134 tun má USA, v Evropě je to Německo, Itálie či Francie. Jde hlavně o „dědictví“ z dob zlatého standardu. Centrální banky tak mohou intenzivně působit na trh se zlatem, a především na jeho cenu. Důsledkem toho evropské centrální banky v roce 1999 podepsali první dohodu centrálních bank o zlatě, která limituje množství zlata, které mohou smluvní strany v jednom roce společně prodat. Od tehdy byly podepsány dvě další dohody v letech 2004 a 2009. V roce 2017 byla celková nabídka 9050 tun zlata, z čehož 72,8 % tvořilo zlato vytěžené z dolů, zbytek je zlato získané z recyklace. Zlato se těží na každém kontinentu s výjimkou Antarktidy, kde je jeho těžba zakázána Madridským protokolem z roku 1991, který ji prohlásil za rezervaci určenou míru a vědě. [18] [21]

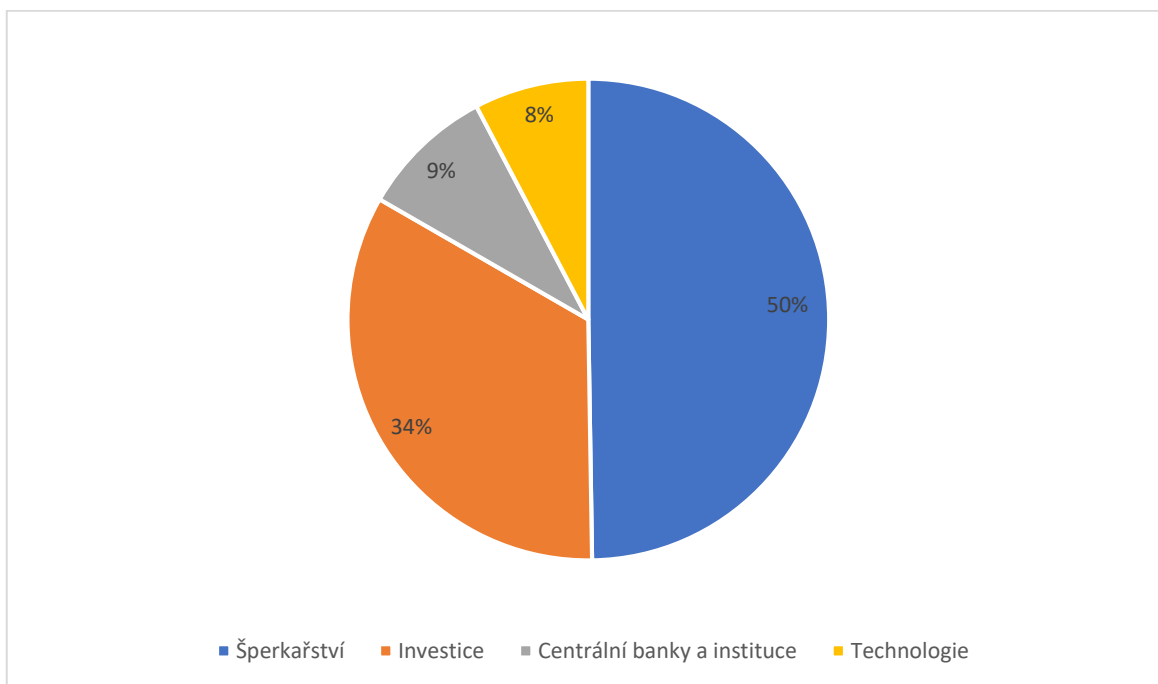
V produkci zlata dominovala od roku 1905 Jihoafrická republika. Podle odhadů více než 50 % veškerého dosud vytěženého zlata tvoří produkce právě z této země. V roce 2007 ji překonala Čína. Od té doby roční produkce zlata v JAR klesají a v minulém roce tak byla až na šestém místě. Pokles produkce v JAR souvisí s vyčerpáním nejpřístupnějších ložisek. Těžbu je tak nutné soustředit více do hloubek, což je samozřejmě ekonomicky nákladnější. Dalším problémem je nespokojenost zaměstnanců, kteří pracují v těžkých a nebezpečných podmínkách. Zvýšení bezpečnosti v dolech rovněž navýšilo náklady a zároveň omezilo těžbu. Čína prvenství v produkci zlata obhájila i loni s procentním podílem na světové

produkci 5,15 %, za níž následuje Austrálie s přibližně 3% podílem. V Evropě se nejvíce zlata těží v Rusku, které je ve světové produkci na třetím místě. Největšími americkými producenty zlata jsou USA, Peru, Kanada, Mexiko a Brazílie. Faktem je, že největší zlaté doly po celém světě spravují a vlastní velké těžební společnosti. Největší je kanadská společnost Barrick Gold, která vlastní doly nejen v Americe ale také v Austrálii, Papua nové Guineji, Zambii a Saúdské Arábii. V roce 2017 vyprodukovala ve všech svých dolech přibližně 5,3 milionů uncí zlata. [22]

3.1.4.2 Poptávka zlata

Graf číslo 4 zobrazuje různé segmenty spotřeby zlata ve světě v roce 2017. Celková poptávka činila asi 8528 tun zlata.

Graf 3 - Poptávka po zlatu v roce 2017 podle odvětví



Zdroj: World Gold Council, 2018 [22]

Jak bylo již zmíněno v úvodu této podkapitoly, zlato se využívá pro různé účely. Dominantní postavení má v oblasti poptávky šperkařství. Spotřeba zlata v tomto segmentu činila ročně téměř 50 % celkové spotřeby. V oblasti výroby šperků jsou největší trhy v Indii a Číně. Poptávka z těchto zemí tvoří přibližně polovinu z celkové poptávky zlata

ve šperkařství. Zlato díky svým jedinečným vlastnostem přispívá investorům k výkonnosti jejich portfolií. Diverzifikuje riziko při negativním vývoji ekonomiky bez snížení dlouhodobých výnosů. V současnosti tvoří zlato spotřebované na investiční účely přibližně třetinu celkové poptávky, který se skládá buď z nepřímého vlastnictví prostřednictvím ETF fondů, nebo z přímého v podobě zlatých mincí a cihel. Poptávka zlata v oblasti investování v roce 2016 rapidně vzrostl oproti předešlému roku. Když v roce 2015 byla spotřeba investovaného zlata přibližně 918,7 tun, v roce 2016 vzrostla až o 69 % na 1561,07 tun. Šlo o nejvyšší hranici od roku 2012. Spotřeba zlata v průmyslu, zejména v elektronickém, díky jeho odolnosti a elektrické vodivosti představuje asi 8 % procent celkové poptávky. Jako jeden z mála prostředků je zlato všeobecně přijímané investičními směrnicemi centrálních bank na celém světě. Od roku 2010 jsou centrální banky čistými kupci zlata. Může za to zpomalení prodeje zlata Evropských centrálních bank spojen se zvýšením poptávky ze strany rozvíjejících se trhů v Latinské Americe, Středním Východě a Asii. Poptávka zlata centrálními bankami v roce 2016 tvořil přibližně 9 % z celkové poptávky. [21]

3.2 Komoditní burza

Obchod s komoditami tzv. trading, probíhá stejně, jako obchod s dalšími podkladovými aktivy na burzovních trzích. Burza je organizovaný trh, což znamená že její fungování je vymezeno zákony, předpisy a pravidly, které je nezbytné striktně dodržovat. Ve většině zemí je nutné k organizování burzovních obchodů získat licenci od příslušného orgánu. [12]

Existují různé druhy burz, rozlišujeme především:

- Peněžní burzy
- Komoditní burzy
- Měnové burzy

Komoditní obchodování probíhá intuitivně na komoditních burzách, proto se jim budeme v následujícím textu blíže věnovat.

Komoditní burza je místo, na kterém se pod striktním dohledem kontrolních orgánů uskutečňují jednotlivé obchody mezi nakupujícími a prodávajícími. [12] V dnešní době

většina obchodů probíhá virtuálně k čemuž přispěl rozvoj informačních technologií. Umožnil, že obchodovat na burze tak dnes může prakticky kdokoliv. Jediné, co potřebuje je přístup na internet. Vytváření cen komodit na základě střetu nabídky s poptávkou, je hlavní funkcí každé komoditní burzy. Ceny, za které se uskutečňují obchody na komoditních burzách odrážejí tržní ceny. Důvodem je jejich rychlé reagování na měnící se podmínky a světový význam, který se jim přikládá. Komoditní burzy díky své významnosti existují ve všech vyspělých zemích světa, ale je důležité podotknout, že se v žádném státě neobchodují všechny světové komodity. Některé mezinárodní komoditní burzy však mají tak velký význam, že se jejich kotacemi se řídí většina OTC obchodů, a to nejen doma, ale i v zahraničí. [1] Nejznámější komoditní burzou na světě je The Chicago Board of Trade (CBOT), která vznikla už v roce 1948. Významné jsou i další jako Chicago Mercantile Exchange (CME), New York Board of Trade (NYBOT). V Evropě můžeme za nejvýznamnější považovat London International Financial Futures Exchange (LIFFE).

Na komoditních burzách se v porovnání s peněžními obchodují aktiva, které mají reálný podklad. Komodity se zde obchodují v tzv. kontraktech, respektive lotech. Každý kontrakt má své specifikované množství základní komodity, a to je vždy určitý násobek ceny, za kterou je obchod uskutečněn. Kontrakty jsou vystaveny na určité konkrétní datum, kdy dojde k jejich vypršení a fyzickému dodání vlastníkově kontraktní listiny.

3.3 Analýza

Obchodník vstupující na trh má možnost se rozhodnout, jakým způsobem chce směřovat svoji obchodní strategii a jaké zdroje k tomu využije. V následujících třech podkapitolách budou představeny základní tři typy analýz.

3.3.1 Psychologická analýza

Psychologická analýza vychází z předpokladu, že burzovní trhy jsou pod silným vlivem masové psychologie burzovního publika, které tím, že působí na účastníky trhu a jejich řízení, ovlivňuje úroveň kurzů. Znamená to, že budoucí vývoj kurzů závisí na impulsech, které ovlivňují chování davu, který se pak přiklání k nákupům nebo prodejem. Za nejvýznamnější teoretické koncepce zabývající se vlivem masové psychologie na kurzy, se považují: [1]

- Keynesovými spekulativní rovnovážná hypotéza
- Kostolanyho burzovní psychologie
- Teorie spekulativních bublin
- Drašnarová koncepce psychologické analýzy

Podle **Keynese** není důležité, jaká bude vnitřní hodnota aktiva, ale jeho tržní cena. To, na čem opravdu záleží, je sentiment investorů, tedy výsledek masové psychologie velkého počtu neznalých individuí. Keynesovská spekulativní rovnovážná hypotéza vychází z toho, že neustále narůstá podíl obchodovaného aktiva v rukou nezkušených investorů oproti kvalifikovaným investorům, a to má za následek, že trhy reagují často nepřiměřeně i na bezvýznamné události a v důsledku toho se stávají příliš volatilními, zvyšuje vliv davu a snaha dosáhnout zisk v krátkém období, proto je důležitější odhad budoucího chování investičního publika než fundamenty. [1]

Kostolany vychází z předpokladu, že v krátkém období, do jednoho roku, jsou kurzy ovlivňovány zejména reakcemi účastníků burzy na různé události, zatímco ve střednědobém a dlouhodobém časovém horizontu mají na vývoj kurzu vliv zejména fundamentální faktory. Kostolany dělí účastníky obchodování na burzách na hráče, kteří tvoří až 90 % burzovního publika, chtějí dosáhnout rychlé zisky, byť menší, rozhodují se emocionálně a jdou vždy s proudem a spekulanty, kteří realizují dlouhodobější transakce, jejich chování není emocionální, vychází z fundamentální analýzy. Obvykle jsou při obchodování úspěšnější než hráči. Základem jeho teorie je zkoumání poměru, v jakém je rozděleno vlastnictví obchodovaného aktiva mezi hráči a spekulanty. Z tohoto poměru odvozuje pohyb kurzu (růst / pokles) a změny objemu obchodu (nárůst / zmenšování). [1]

Teorie spekulativních bublin hovoří o vzniku situací, kdy kurz určitou dobu roste (nebo klesá) bez jakéhokoli rozumného vysvětlení podloženého fundamentálním vysvětlením. Tento pohyb se po určité době náhle a neočekávaně zastaví a kurz se velmi prudce pohne najednou opačným směrem. Menší, nezkušení investoři při vzniku těchto bublin utrpí často velké ztráty. [1]

Drašnarová koncepce psychologické analýzy vychází ze dvou vlastností lidí. Z chamtivosti a strachu. Když převládá chamtivost, lidé nakupují, poptávka a kurzy rostou. Když se však dostaví strach ze ztráty nabytého bohatství, lidé se leknou a začnou prodávat,

nabídka převyšší poptávku a kurz začne klesat. Ze strachu může vzniknout až panika, a pak dochází ke skokovému poklesu hodnoty. [1]

3.3.2 Fundamentální analýza

I když se fundamentální analýza (dále jen FA) využívá zejména při obchodování s akcemi, můžeme v definici fundamentální akciové analýzy podle Rejnuše pojem akce nahradit širším pojmem aktivum a říct, že fundamentální analýza je založena na předpokladu, že vnitřní hodnoty (teoretické ceny) aktiv se liší od jejich aktuálních tržních cen (kurzů), za které jsou obchodovány na veřejných (organizovaných) trzích. FA zkoumá vztah vnitřní a tržní ceny. Pokud je vnitřní hodnota aktiva vyšší než jeho kurz, je aktivum považováno za podhodnocené, a pokud je naopak ve srovnání s kurzem vnitřní hodnota aktiva nižší, tak je aktivum nadhodnoceno. FA hledá a analyzuje faktory, které ovlivňují vnitřní hodnotu aktiva. FA kromě výpočtu vnitřní hodnoty aktiva dokáže poskytnout krátkodobé i dlouhodobé projekce ekonomické a finanční síly země, odvětví, sektoru a společností, příp. určitého aktiva, především akcií. Fundamentální analýzu, resp. některé její části, využívají obchodníci s akcemi, komoditami i jmény. Z hlediska charakteru zkoumaných faktorů lze rozdělit FA na tři typy: [1] [13]

1. **Globální (makroekonomická) analýza** analyzuje ekonomiku jako celek a zkoumá vztahy mezi vývojem tzv. globálních, především makroekonomických agregátů a pohybem kurzů.
2. **Odvětvová analýza** analyzuje specifika jednotlivých konkrétních odvětví a následně předpovídá perspektivy jejich budoucího vývoje.
3. **Analýza jednotlivých společností** (mikroekonomické) analýza zkoumá vnitřní parametry konkrétních akciových společností v kontextu s vlastnostmi akcí, které tyto společnosti emitují, a to z pohledu retrospektivy, současnosti a perspektivy – výhledu do budoucnosti. [1] [13]

3.3.3 Technická analýza

Technická analýza je nástroj na předpověď budoucího vývoje hodnot kurzů komodit, měn, akcií a indexů. Je to „studium pohybu ceny.“ [23] Na základě historických dat se snaží předpovědět pohyby cen/kurzů v budoucnosti. Vychází z publikovaných tržních dat, kterými

jsou tržní kurzy, hodnoty indexů, a dokonce i objemy zrealizovaných obchodů. Zastánci technické analýzy obecně „vycházejí z přesvědčení, že rozhodujícími faktory, z nichž je možné při analýze vycházet, jsou tržní nabídka a poptávka, na základě jejichž střetu je na organizovaných kapitálových trzích tvořen kurz. V něm je, podle jejich názoru, už dostatečně obsažený optimismus a pesimismus všech účastníků trhu, ale také všechny jimi uvažované fundamentální údaje a další informace na příslušném trhu dostupné. Kromě těchto údajů vycházejí i z předpokladu, že lidské chování zůstává v podstatě neustále stejné, a na základě tohoto názoru ukazují, že stejně i chování investorů se vyznačuje opakujícími se reakcemi. Z těchto reakcí následně odhadují budoucí vývoj kurzů. Jelikož historie cenových změn se v průběhu času neustále opakuje, pomocí časových řad identifikují příslušné vývojové trendy a cenové rámce.“ [24] „Cílem technické analýzy je tedy zachytit formou grafického znázornění vývoj kurzů akcií a burzovních indexů, a následnými rozbory vytvořených grafů zjišťovat budoucí směry jejich dalšího vývoje.“ [24] Technickou analýzu lze, podle nástrojů, které využívá, lze rozdělit na grafickou analýzu a analýzu založenou na využití technických indikátorů. Grafická analýza využívá především grafy a údaje v nich zakořeněné, které jsou viditelné pouhým okem. Technická analýza využívající technických indikátorů je založena právě na zkoumání vnitřních souvislostí v grafech. K tomuto využívá tzv. trendové indikátory, oscilátory a smíšené indikátory, které jsou kombinací trendových ukazatelů a oscilátorů. Smíšenými ukazateli jsou například New High - New Low Index a Put-Call Ratio, které mohou jít shodně s trendem nebo dokonce i před ním. „Umožňují náhled do intenzity býčích nebo medvědích projevů na trhu.“ [25] Trendové indikátory, oscilátory a prediktivní indikátory budou popsány níže.

3.4 Indikátory technické analýzy

Technické indikátory jsou různé matematické funkce, vzorce, výpočty, kterými se dostávají informace zejména o tom, že je některé aktivum překoupené (došlo k příliš rychlému vzrůstu ceny a ta bude proto s velkou pravděpodobností korigovaná směrem dolů) nebo přeprodané (po prudkém poklesu ceny). Také nám poskytují informace o tom, že kurz změnil nebo brzy pravděpodobně změni směr a může to být začátek nového silného trendu – vzrůstu nebo poklesu.

Technické indikátory využívají různé faktory, ale vycházejí z vývoje kurzů a objemů obchodu v čase. Slouží pro určení analýzy budoucího vývoje kurzů, včetně možné indikace nákupních a prodejních signálů, a také k predikci budoucího vývoje celých trhů. K základním druhům technických indikátorů patří: [1]

- Trendové indikátory
- Oscilátory
- Prediktivní indikátory

3.4.1 Indikátory

Z grafů mohou být některé informace viditelné na první pohled, ale množství informací vyjde na povrch až po důkladné analýze pohybů na grafu. Analýza zkoumající tyto pohyby a jejich podstatu se nazývá technická analýza. Ta využívá technických indikátorů pro analyzování trhu a následné rozhodování se pro vstup/výstup z obchodu. Jedněmi z těchto technických ukazatelů jsou trendové indikátory. Ty, jak již název naznačuje, jsou založeny právě na analýze současného trendu. Trendové ukazatele jsou opožděné za trendem a jejich vývoj se obrátí, jakmile dojde ke změně trendu. [25] Zpoždění za trendem může být větší nebo menší. Důležité však je, že jejich změna nastává vždy v závislosti na změně trendu. Tedy pokud se změní trend, změní se i trendový indikátor. Zařazují se sem například klouzavé průměry, MACD a další.

Jedněmi z nejzákladnějších a nejjednodušších trendových indikátorů jsou **klouzavé průměry**. Klouzavé průměry se označují zkratkou MA, pocházející z angličtiny, což znamená Moving Average. Klouzavý průměr je průměrem několika posledních údajů, resp. průměrnou hodnotou hodnot naměřených za zvolený časový úsek. Zvolený počet údajů, resp. délka časového úseku se obvykle označuje n a je libovolně volitelná, to znamená, že je možné si zvolit z jakého velkého výběru (časového rámce) bude klouzavý průměr počítán. Například MA_3 je klouzavý průměr vypočítaný za poslední 3 časové úseky (3 dny, 3 hodiny, ... - záleží podle zvoleného grafu). Když spojíme hodnoty všech vypočtených klouzavých průměrů, dostaneme graf zvoleného klouzavého průměru. Klouzavý průměr se nazývá klouzavým proto, že s novou hodnotou na grafu se znovu vypočítá průměr z posledních n hodnot, a to tak, že se nejstarší hodnota odstraní z výpočtu a nahradí nejnovější. Zvolený

časový rámec se dá v podstatě představit jako nějaké okno, které se klouže po grafu a ukazuje hodnoty, z nichž je aktuální klouzavý průměr počítán. [26]

Existují tři základní druhy klouzavých průměrů:

- Jednoduchý (MA)
- Exponenciální (EMA)
- Vážený (WMA)

Výpočet jednoduchého klouzavého průměru je totožný s výpočtem aritmetického průměru. **Exponenciální klouzavý průměr** je vhodnější pro použití jako jednoduchý MA, neboť přikládá větší váhu novějším hodnotám. To je z praktického hlediska opravdu vhodné, neboť čím jsou hodnoty starší, tím jsou pro nás méně zajímavé, a nejnovější hodnoty pro nás mají zpravidla největší hodnotu, neboť jsou nositeli nejaktuálnějších informací. Posledním ze zmíněných typů klouzavých průměrů je **vážený průměr**. Jelikož však vážený průměr může být silně individualistickou záležitostí a může nabývat nejrůznějších podob, tento nebude velmi rozebírán. Při jeho konstrukci se časovému úseku (konkrétním hodnotám), zpravidla na základě dlouhodobého zkoumání, přiřadí na základě pozorování přesně spočítané hladiny významnosti.

Z trendových ukazatelů bych blíže přiblížil ještě alespoň jeden, a to **MACD-Histogram**. MACD nebo Moving Average Convergence-Divergence je indikátor zkonstruován analytikem Geraldem Appelem. V překladu název tohoto ukazatele znamená sbíhavost – rozbíhavost klouzavých průměrů. MACD je systém exponenciálních klouzavých průměrů. Je složen z kombinace tří klouzavých průměrů. Rozdíl dvou z nich tvoří jednu křivku – bývá zobrazována jako plná čára a nazývá se MACD. Druhá křivka je exponenciálním klouzavým průměrem první křivky a je nazývána jako signální křivka (Signal). [25]

3.4.2 Oscilátory

Oscilátory jsou ukazatele, které „pomáhají identifikovat body zvratu“ [25] ve vývoji cen, tedy body, kdy dochází ke změně vývoje trendu. Mohou jít shodně s trendem nebo dokonce i před ním. Často se jejich průběh obrátí dříve než samotný trend, čímž dávají dobré

signály pro nákup, resp. prodej. Z oscilátorů je určitě nutné zmínit Stochastic, Momentum a RSI.

Stochastic je ukazatelem hybnosti vyvinutým Georgeem C. Laneem v 50. letech 20. století, který ukazuje pozici poslední závěrečné ceny vzhledem k předchozímu vysokému a nízkému rozmezí. Ukazatel měří hybnost porovnáním závěrečné ceny s předchozím obchodovacím rozsahem za určité časové období. Stochastický ukazatel nesleduje cenu nebo objem podkladového měnového páru, ale rychlost a hybnost ceny. To znamená, že stochastický ukazatel mění směr před samotnou cenou a lze jej tedy považovat za hlavní ukazatel. Nejdůležitější signály, které Lane identifikoval, jsou býčí a medvědí divergence, které se vytvářejí na stochastickém ukazateli, což může předvídat nadcházející cenové zvraty. Protože však stochastický indikátor kolísá v určitém rozmezí, lze jej také použít k identifikaci překoupených a přeprodaných cenových hladin. Indikátor **Momentum** pracuje na podobném principu. Účelem **RSI** je zmapovat současnou a historickou sílu nebo slabost finančních instrumentů nebo trhu na základě uzavírací ceny posledního obchodovacího období. Nejčastěji se používá s nastavením posledních 14 období a rovněž se zobrazuje na stupnici 0 až 100 bodů. Za vysoké a nízké hladiny se považují hodnoty větší než 70 a menší než 30. [25]

3.4.3 Prediktivní indikátory

Toto je speciální skupina indikátorů, které přímo s určitou pravděpodobností odhadují směr budoucího pohybu trhu. Rozdíl oproti trendovým indikátorům, či oscilátorům je ten, že prediktivní indikátory se vykreslují do budoucích časových úseků. Jedná se o speciální skupinu indikátorů, která není volně dostupná v platformě MetaTrader4, a proto se musí doprogramovat. Prediktivní indikátor je například námi vytvořený ARIMA/GARCH indikátor, který vykresluje budoucí hodnoty predikce do grafu.

3.5 Risk management

Risk management je zejména řízení podstupovaného rizika a jednu pozici. Riziko se tak vypočítá a zafixuje tzv. stop lossem. Stop loss říká, při jaké maximální možné ztrátě již se nevyplatí pozici dále držet.

Velikost stop lossu je možné měnit v závislosti na aktuální volatilitě na trhu. Při zvýšené volatilitě jsou trhy více rozkolísané, proto je efektivně nastavovat stop lossy na vyšší částky. Pokryje se tím nechtěné vyhození z obchodu. Pokud je naopak volatilita na daném trhu nízká, znamená to, že je možné velikost stop lossu zmenšit. Optimalizování velikosti stop lossu se považuje za pokročilou techniku v obchodování a přináší o to větší zisky. Obchodník je tak schopen získat maximální možný zisk z daného obchodu a minimalizovat ztráty. [27]

3.5.1 Risk-Reward-Ratio (RRR)

Obchodní strategie by měla mít kromě vstupů a výstupů jasně danou i úroveň RRR, což je zkratka pro risk-reward-ratio. Jedná se o poměr zisku k riziku. Tj. poměr sumy kterou obchodník riskuje na jeden obchod a částky, kterou chce v obchodě získat. Čím větší RRR tím, menší má zpravidla strategie úspěšnost. Pokud obchodník například používá RRR 1: 2 tak, aby byl v celkovém pohledu v zisku potřebuje mít alespoň cca 40 % úspěšnost obchodů. V tomto případě 6 z deseti obchodů skončí stop-lossem a tedy ztrátou a 4 obchody skončí ziskem. V konečném důsledku obchodník vydělá. Při použití RRR 1: 1 musí obchodník dosahovat alespoň 51 % úspěšnost, lépe je však začít v takovém případě na 60 %. Larry Williams známý burzovní obchodník dosahoval úspěšnost strategie nad 80 procent ovšem za cenu negativního RRR (to znamená, že více riskoval na jeden obchod, než mohl získat). [12]

3.6 Obchodní systémy

Obchodní systémy jsou předem nastavená pravidla, dle nich se nakupuje a prodává. Obchodní systém by měl odrážet psychiku jedince, který dle daného obchodu obchoduje. Jednoduše řečeno by mu měl vyhovovat. Obchodní systém může využívat čistě cenových pohybů (price action), případně využívat indikátorů signalizující potenciální obchod.

Dle náročnosti obchodování můžeme rozdělit do několika skupin:

- **Investování** je z hlediska náročnosti nejjednodušší strategie Buy and Hold. Je založená na fundamentální analýze a dané pozice jsou drženy v řádech měsíců až let. Nejznámějším představitelem této strategie je Warren Buffett.

- **Poziční obchodování** je založeno také na fundamentální analýze, ale s technickými cenovými hladinami k nákupu či prodeji. Dané pozice pak jsou drženy v rádech týdnů až měsíců.
- **Swingové obchodování** je založeno primárně na technické analýze. Dané pozice jsou drženy několik málo hodin, či dní. Tato forma strategie je považována za statisticky nejziskovější strategií v obchodování.
- **Intradenní obchodování**, někdy označované jako day trading, či intraday trading je stylem, kdy žádná pozice nezůstává přes noc, a tak každý den defacto začíná obchodník od nuly. Využívají se k tomu krátké časové rámce 1 – 30 minut. Většina signálů je založena čistě na technické analýze.
- **Scalping** je obchodní styl, kde se pozice drží velmi krátkou dobu v rádech vteřin, maximálně minut. Jedná se o strategii, která je velice zisková pro brokera, jelikož obchodník platí poplatek za každý uskutečněný obchod, kterých je velké množství. Potenciálně je velmi zisková pro obchodníka, avšak s nejvyšším možným rizikem, protože v případě neúspěšného obchodu se jedná o velkou ztrátu. [23]

Pro potřeby našeho systému budeme obchodovat intradenně, ale ke vstupům budeme používat jednohodinové grafy, což je typické pro swingové obchodování. Obchody budeme následně držet přibližně 12 hodin.

Obchodní systémy se dále mohou dělit podle úrovně své „inteligence“ umět signalizovat obchod, či jejich schopnosti automatizace. Z tohoto pohledu můžeme obchodní systémy rozřadit do třech kategorií. [28]

1. **Manuální obchodní systémy** – obchod je zde vyhodnocován i zadáván ručně.
2. **Poloautomatické obchodní systémy** – systém obchody neuzavírá sám, ale skenuje trh a odešle upozornění o blížící se obchodní příležitosti.
3. **Plně automatické obchodní systémy (AOS)** – systém je schopný vyhodnotit situaci na trhu samostatně a dle toho zadává příkazy k otevírání a uzavírání obchodů.

Při vytváření AOS se postupuje od manuálních systémů výš. Pakliže se v manuálním systému jedna a tatáž situace opakuje a má stálý vzorec na vyhodnocení, je možné tuto úlohu předat obchodnímu systému. Vznikne tak poloautomatický obchodní systém, který funguje

a kterému pouze přidáváme pravomoci k řízení obchodních pozic. V okamžiku, kdy systém na základě předchozích zkušeností funguje samostatně, můžeme hovořit o automatizovaném obchodním systému. Další variantou je vytvoření algoritmu a následné backtestování na historických datech. Zde ovšem vzniká zkreslení z toho důvodu, že se trh neustále mění a to, že systém fungoval na historických datech, nemusí nutně znamenat, že bude fungovat na datech budoucích. Výhoda takto vytvářeného systému je rychlost jeho vytvoření a dá nám alespoň základní informace o tom, zda by naše myšlenka automatického obchodního systému mohla být správná. [29]

3.6.1 Typy obchodních platforem

Na výběru obchodní platformy by mělo hrát významnou roli rozhodnutí, pro jaký trh se obchodník rozhoduje. Ne každá platforma totiž poskytuje možnosti obchodů s požadovanými komoditami. Dalším významným aspektem je také výběr brokera. Broker totiž často poskytuje své vlastní platformy, a proto je třeba zvolit platformu, na kterou je možné připojit i jiné brokery. Na první pohled se to zdá jako drobnost, ale změna platformy kvůli změně brokera může zabrat nějaký čas na učení s novým prostředím.

Platformu je třeba vybírat i pro její analytické funkce, jako jsou například různé typy indikátorů, oscilátorů, grafů, dle kterých se provádí většina rozhodnutí v technické analýze. Intradenní obchodování se zase neobejde bez možnosti naprogramování vlastního kódu, pomocí kterého zjednodušíme rozhodování o provedení, či neprovedení obchodu. Není to vždy zárukou dlouhodobého zisku. Je třeba často do kódu zasahovat a optimalizovat ho. [30]

Příklady platforem:

- MetaTrader
- Plus500
- AmiBroker
- Sirix webtrader
- Ninja Trader

3.6.2 MetaTrader 4

Pro vytvoření obchodního systému byl vybrán program MetaTrader 4 (MT4). Tento program disponuje programovacím jazykem MQL4 (vycházejícího z jazyka C), který bude použit pro vytvoření prediktivního indikátoru ARIMA/GARCH, na jehož základě bude obchodování probíhat. Podstatné je, že jazyk MQL4 lze propojit s jazykem R pro statistickou analýzu, jelikož surové MQL4 pokročilé matematicko-statistické metody pro predikce časových řad neovládá. Další podstatnou výhodou platformy MT4 je funkce optimalizace, jenž využívá neuronových sítí a genetických algoritmů. [29] Přesný manuál pro jednotlivé funkce je k dispozici na oficiálních stránkách s podporou na www.mql4.com. Testování na reálném trhu bude prováděno z tzv. demo účtu, který umožňuje disponovat fiktivními 100 000\$. Tento účet byl vytvořen u brokerské společnosti na stránkách www.xtb.com.

3.6.3 Rizika online obchodování

Tato podkapitola se zabývá riziky, která mohou vznikat nebo vznikají při online obchodování na komoditních burzách.

3.6.3.1 Latence

Zpoždění představuje v oblasti real-time nakupování značný problém. V případě, že obchodník chce provést kontrakt, ale má vysokou latenci, může se stát, že se jeho kontrakt provede v hodnotách pro něj nepříznivých. Tímto způsobem může obchodníkovi vznikat ztráta jen díky špatnému připojení. Je důležité, aby měl obchodník stabilní připojení a kvalitní softwarové a hardwarové prostředky. V případě výpadku by měl mít obchodník i záložní přístup k internetu, např. mobilní hot-spot, restauraci s bezdrátovým připojením apod. [29][30]

3.6.3.2 Lidský faktor

Lidský faktor představuje v obchodování poměrně velké riziko. Spousta nových obchodníků skončí právě na chybě lidského faktoru. Aby bylo možné předejít chybě lidského faktoru, tak je zapotřebí, aby si obchodník vytvořil strategii. Tuto strategii by si měl nejprve otestovat na starších datech a virtuálních penězích za pomoci demo účtu. Pokud se strategie osvědčí, je nutné ji po nasazení striktně dodržovat. [29][30]

3.6.3.3 Obchodování s komoditami

Obchodování s komoditami samo o sobě představuje poměrně velké riziko, kde v případě neuváženého neúspěchu přijde obchodník o vložený kapitál. V případě, že hlavním zdrojem obživy obchodníka je právě obchodování, je toto riziko možné brát jako existenční. [12]

3.6.3.4 Problém s hardware

Jediným možným prostředkem při online obchodování je používání určitého HW (počítač, tablet, mobilní telefon), který ale může mít poruchu. V průběhu poruchy není obchodník schopen zadávat příkazy, nebo měnit obchodní strategie, protože není možné zadat do obchodního systému své pokyny. Toto může obchodníka připravit o jeho zisk. V případě vybudování vlastního automatického obchodního systému je možné toto riziko snížit, protože veškeré příkazy se vykonávají automaticky, a to na HW konkrétního brokera, pokud tuto službu podporuje. Toto riziko se ovšem nedá vyloučit, protože i na straně brokera může dojít k poruše HW vybavení. Pro snížení problémů, které nastanou při problémech s HW, je vhodné mít k dispozici více než jeden kus HW, který obchodník bude moci použít. Například pokud obchodníkovi vypoví službu počítač, je dobré mít v záloze například tablet, na kterém obchodník může dále provádět své příkazy. [29]

3.6.4 Backtest

Backtest se provádí na historických datech a poskytuje obchodníkovi představu o tom, jak se jeho strategii dařilo v minulosti. Podle něj se dostanou relativně přesné informace o tom, kolik ztrátových obchodů za sebou se může očekávat, jak vysoký stop loss optimálně zvolit a zda je strategie dostatečně robustní. Tj. zda není přeoptimalizovaná a nemá v sobě příliš mnoho parametrů, které fungují pouze za určitých podmínek. Trhy se neustále mění, a tak právě robustnost a jednoduchost strategie je to o co by se měl každý obchodník snažit. Vynikající pomůcka v backtestování je analýza MAE a MFE. Jedná se o hodnoty, které představují, jaký maximální stop loss má obchodník zvolit, aby ho udržel v dobré pozici a jaký zisk je schopen očekávat na jeden obchod. Z toho vyplývá, že je lepší vzít otevřený profit než čekat na vyšší zisk, protože statisticky máme prokázáno, že ve většině případů to bude zbytečné. Backtest můžeme rozdělit na ruční nebo prováděný počítačem.

Ruční backtest – představuje ruční procházení grafů a vyhledávání vstupů a výstupů

Počítačový backtest – spočívá v zadání vstupních a výstupních parametrů a počítač udělá v podstatě veškerou práci za nás, přičemž nám na konci sdělí parametry strategie s konečným výsledkem. [29]

4. Vlastní práce

4.1 Analýza v R

V této kapitole je ukázáno, jak využít teoretické znalosti z předchozích částí k prediktivnímu modelování časových řad a k vytvoření backtestu strategie na historických datech. Všechny postupy budeme realizovat pomocí programu R, který je volně dostupný na internetové adrese www.rproject.org. V úvodu se v stručnosti seznámíme s balíky a funkcemi, které budou při práci s využitím programu R využívány. V dalších částech budeme pracovat s konkrétními daty, použité postupy budeme ilustrovat graficky, přičemž zdrojový kód v R je k dispozici v přílohách na konci této práce.

4.1.1 Použité funkce v programu R

V této části se seznámíme se základními funkcemi z balíčků obsažených v programu R, které budeme v dalších částech využívat. Půjde však o stručný popis vybraných funkcí, které se vyskytují v použitých zdrojových kódech. Pokud by měl čtenář zájem o úplný seznam funkcí v daném balíku, najde ho v nápovědě k balíku, která se zobrazuje příkazem `?název_balíku`. Podrobné a užitečné informace o jednotlivých funkcích jsou stejně tak k dispozici v nápovědě, která se obdobně vyvolá příkazem `?název_funkce`. Ne všechny zmíněné balíčky jsou součástí standardní instalace programu R, ovšem dají se do něj jednoduše doinstalovat pomocí menu Packages -> Install packages. [31]

4.1.1.1 Balíček Base

`diff(x)` – Tato funkce volána na časovou řadu $\{X_t\}$ vrátí jednu diferencovanou časovou řadu $\{X_t\}$. Pro časovou řadu $\{X_t\} = \{X_1, X_2, \dots, X_T\}$ tedy vrátí diferencovanou časovou řadu $\{\Delta X_T\} = \{X_2 - X_1, X_3 - X_2, \dots, X_T - X_{T-1}\}$.

`sumsum(x)` – Volání této funkce na časovou řadu $\{X_t\}$ vrátí její kumulovaný součet. V našem případě kumulovaný součet zisku z jednotlivých obchodů při backtestu.

`as.data.frame(x)` – Volání této funkce na vektor, matici, či jiný datový formát napřed zjistí, zda je možné převést data do formátu `data.frame`. Pokud ano, vrátí formát `data.frame` se vstupními daty.

4.1.1.2 Balíček Readxl

`Read_excel(x)` – Jako parametr x je cesta ke xls souboru se staženými testovacími daty. Funkce slouží jen pro export dat z Excelu do prostředí R. Formát vrácených dat je datový formát `tbl_df`, ze kterého je možné vytvořit formát `data.frame` pomocí funkce `as.data.frame()`, se kterým pracuje většina funkcí.

4.1.1.3 Balíček Forecast

`arima(x)` – Jak už název funkce napovídá, tato funkce slouží pro výpočet modelu $ARIMA(p,d,q)$. Funkce nabízí řadu parametrů, avšak pro nás jsou nejdůležitějšími parametry vstupní data a vektor $c(p,d,q)$, obsahující parametry $ARIMA$, pro které se model vypočítá.

4.1.1.4 Balíček Xts

`xts(x)` – Tato funkce vytvoří rozšířený objekt časové řady. Zejména jde o přidání datumů k jednotlivým záznamům ve vektoru časové řady.

4.1.1.5 Balíček Stats

`AIC(x)` – Je obecná funkce vypočítávající Akaikeovo informační kritérium pro modelový objekt dle vzorce (6.1). V našem případě vypočítá AIC pro model $ARIMA$.

`Box.test(x)` – Vypočítá statistiku Ljung-Boxova testu na otestování nulové hypotézy, že v časové řadě není přítomna autokorelace. Prakticky díky této metodě testujeme přítomnost, či nepřítomnost autokorelace reziduí.

`lag(x)` – Vypočítá zpožděnou verzi časové řady, respektive časovou řadu posune zpět o určitý počet pozorování.

`residuals(x)` – Je obecná funkce, která extrahuje residua z objektů vrácenými modelovými funkcemi. Tato funkce umožňuje získat standardizovaná residua a to `residuals(model, standardize = TRUE)`, která jsou vhodnější pro otestování autokorelace reziduí.

4.1.1.6 Balíček Rugarch

`ugarchspec(x)` – Tato vytvoří univariační objekt specifikace GARCH modelu před aplikací na časovou řadu. Tato funkce má více parametrů, které je třeba si rozebrat. Prvním zadávaným a nejdůležitějším parametrem je `variance.model = list(garchOrder = c(1,1))`, kterým specifikujeme parametry modelu GARCH(p,q). Další parametr je `mean.model <- list(armaOrder <- c(r,m), include.mean = T)`, který obsahuje střední specifikaci modelu. Jelikož budeme vytvářet smíšený ARIMA/GARCH model, je nutné předat parametry ARMA(r,m). Zároveň si přejeme zahrnout do výpočtu průměr. Třetí používaný parametr je `distribution.model = "sged"`. Ten určuje, jaké má být podmíněné rozdělení odchylek. Vzhledem k použitým časovým řadám byl použit formát "sged", který určuje zobecněné rozdělení odchylek zkosením.

`ugarchfit(x)` – Tato funkce je přímo navazující na předchozí funkci `ugarchspec(x)`, jelikož dříve specifikovaný model vyplní daty za pomoci specifikované iterační metody. Vzhledem k tomu, že model má být co nejvíce autonomní, použili jsme tuto funkci s nastavením `ugarchfit(model, data, solver = 'hybrid')`. Hybridní iterační metoda, totiž postupně zkouší všechny dostupné metody a vybere tu, která nejvíce vyhovuje. Nevýhoda při tomto řešení je výpočetní náročnost procesu.

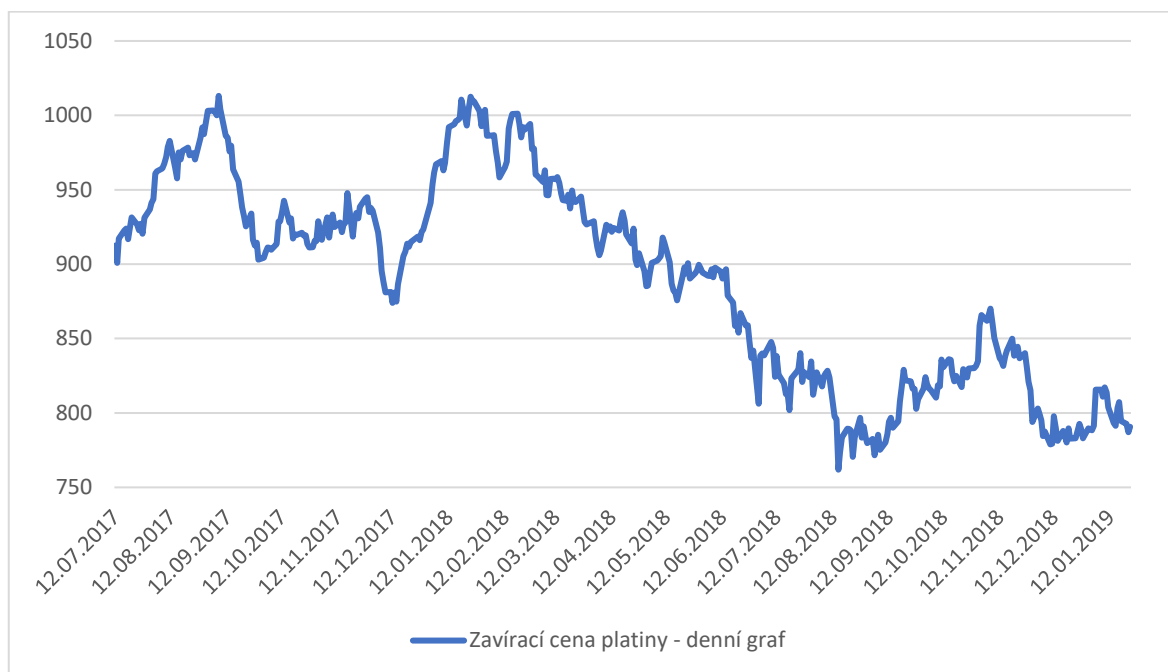
`ugarchboot(x)` – Metoda vytvoří předpověď pro n období na základě vyplněného ARIMA/GARCH modelu pomocí funkce `ugarchfit(x)`

4.1.2 Analýza a predikce v R

Nejprve je třeba vytvořit algoritmus, který bude schopen vytvářet prediktivní modely. Získané poznatky jsme aplikovali na konkrétní data, respektive na platinu v období od 12.7.2017 do 23.1.2019. Celková délka referenčního období tedy čítá 400 záznamů. Historická data jsou stažena z platformy MetaTrader4. Historická data se nestahují od brokera, nýbrž z centrální databáze serveru mql5.com. Server mql5.com uchovává vždy jen omezený počet záznamů. Rozdíl oproti reálným datům zprostředkovaných brokerem může být zpoždění příkazů a nepřesnost přenosu dat a hlavně to, že historická data v sobě nenesou informace o každém jednotlivém pohybu trhu. Z toho důvodu je poté dobré určitý

system testovat na reálných datech. Nicméně pro vytvoření prediktivního modelu, který bude náš systém používat, tato data dostačují.

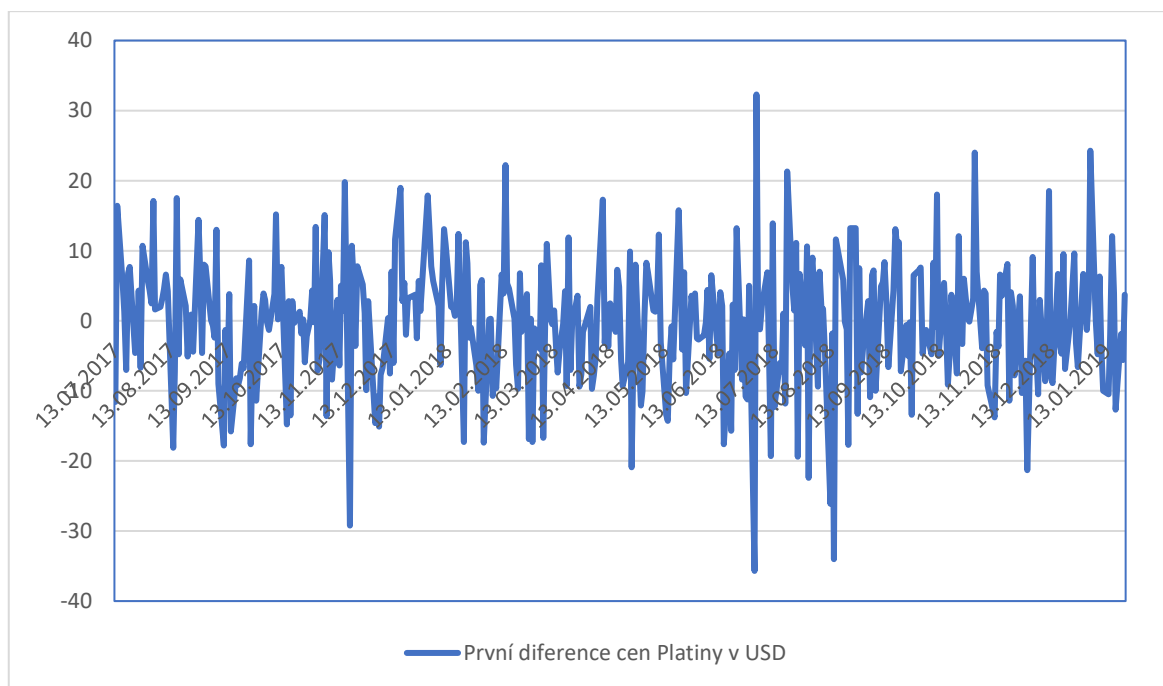
Graf 4 - Zavírací denní ceny platiny v USD



Zdroj: Data stažená z obchodní platformy MetaTrader4 ke dni 13.1.2019, vlastní zpracování

Nejprve je třeba provést identifikaci modelu. To znamená určit, který model bude vyhovovat našim datům. Pro to, abychom mohli model identifikovat, musí být daná řada stacionární. Z grafu 1 je na první pohled zřejmé, že daná řada stacionární není. Proto provedeme diferenciaci prvního řádu, která by mohla danou časovou řadu zestacionarizovat.

Graf 5 - První diference denních cen platiny v USD



Zdroj: Data stažená z obchodní platformy MetaTrader4 ke dni 13.1.2019, vlastní zpracování

První difference časové řady na první pohled vykazuje stacionární průběh. První difference u většiny časových řad stačí k získání stacionární časové řady. Na této časové řadě můžeme provádět odhady modelu a následně prediktivní výpočty.

Postup je takový, že na diferencovanou časovou řadu budeme postupně zkoušet výpočty $ARMA(p,q)$ různých řádů $p = 0, 1, \dots, 5$, a $q = 0, 1, \dots, 5$ s výjimkou $ARMA(0,0)$, jelikož takovýto model nedává ze samotné podstaty smysl a tato podmínka musí být ošetřena v kódu. Maximální řád 5 byl zvolen proto, protože pro modelování $ARMA$ jsou tyto řády nejčastější a je zde mnohem menší riziko zanesení nevýznamného parametru.

Vzhledem k tomu, že se stává, že model často nejde odhadnout jen pomocí autokorelačních funkcí ACF a PACF, přihlíží se k více kritériím. Takové kritérium může být rozptýl, statistická významnost parametrů nebo Akaikeho informační kritérium (AIC). Z důvodů použitelnosti a jednoduchosti jsme se rozhodli právě pro AIC. Nejlepší model dle AIC je ten, který vykazuje nejnižší hodnotu.

Tabulka 3 - Porovnání jednotlivých modelů ARMA(p,q) dle AIC

p	q	AIC
0	1	2895,968
0	2	2894,353
0	3	2894,913
0	4	2894,987
0	5	2896,987
1	0	2895,342
1	1	2896,759
1	2	2895,731
1	3	2894,853
1	4	2896,853
1	5	2898,312
2	0	2895,738
2	1	2896,675
2	2	2895,461
2	3	2896,88
2	4	2898,955
2	5	2900,925
3	0	2894,717
3	1	2896,276
3	2	2897,728
3	3	2898,216
3	4	2899,895
4	0	2895,898
4	1	2897,637
4	2	2897,233
4	3	2899,677
4	4	2899,516
4	5	2901,519
5	0	2897,329
5	1	2899,319
5	2	2899,11
5	3	2901,229
5	4	2901,515
5	5	2900,648

Zdroj: vlastní zpracování

Chybějící výpočty nejsou v tabulce z toho důvodu, že parametry ARMA byly nevýznamné. V případě nestacionární časové řady by se neprovedl výpočet žádný. Z výše uvedené tabulky je patrné, že nejlepší parametry pro model ARMA jsou ARMA(0,2). V tomto konkrétním modelu je autoregresivní složka zcela vypuštěna. Nejvýznamnější složkou je zde proces klouzavých průměrů. MA(2). Přitom neustále platí, že použitá ARMA modelu na prvních diferencích časové řady dává model ARIMA. V tomto případě je model ARIMA(0,1,2).

Tabulka 4 - Odhadnuté parametry modelu ARIMA

Odhadnuté parametry modelu ARIMA		
Call: arima(x = returns, order = (0,0,2))		
	ma1	ma2
Coefficients:	-0.1152	0.1044
s.e.	0.0503	0.0547
sigma^2 estimated as 81.13: log likelihood = -1443.18, aic = 2894.35		

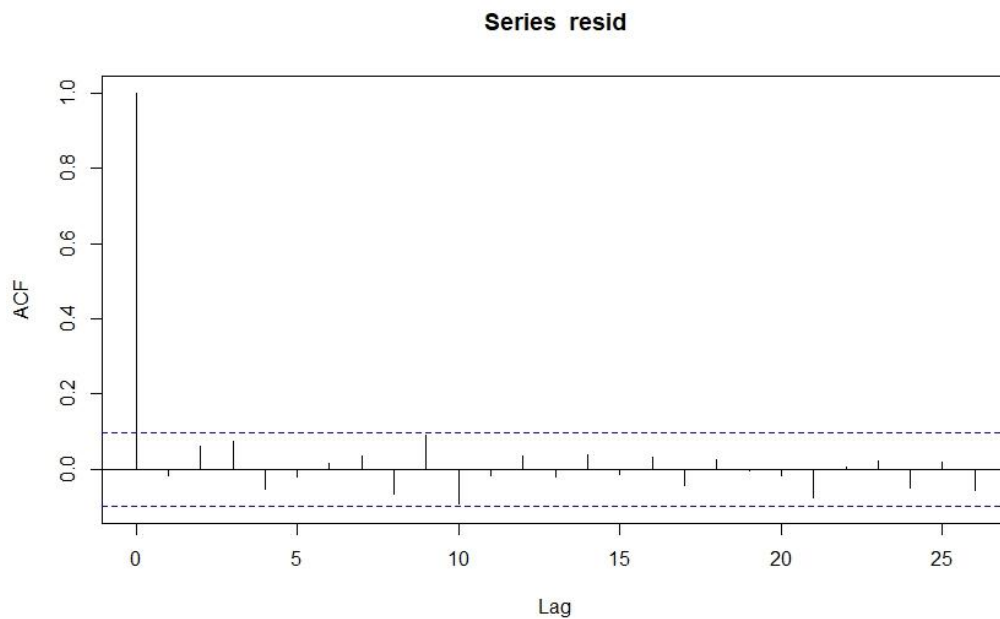
Zdroj: vlastní zpracování – Výstup programu RStudio

Abychom dokázali zachytit podmíněnou závislost reziduí, tak se pro kompletnost modelu na zvolený model aplikuje model GARCH(1,1) pomocí metody `ugarchspec()` a následně `ugarchfit()` pro aplikování dat na zvolený model. Vznikne tak smíšený model ARIMA/GARCH. Pomocí tohoto modelu budeme predikovat budoucí vývoj cen. Nejprve je ale nutné model ověřit zejména na možnou přítomnost autokorelaci reziduí. Pakliže se autokorelace reziduí vyskytuje, není vhodné model používat pro predikce budoucího vývoje.

Přítomnost autokorelace reziduí otestujeme aplikováním Ljung.box testu. P-hodnota Ljung.boxova testu vyšla 0.6260522, což je více než hladina významnosti, takže můžeme potvrdit nulovou hypotézu, že se v modelu nevyskytuje autokorelace reziduí.

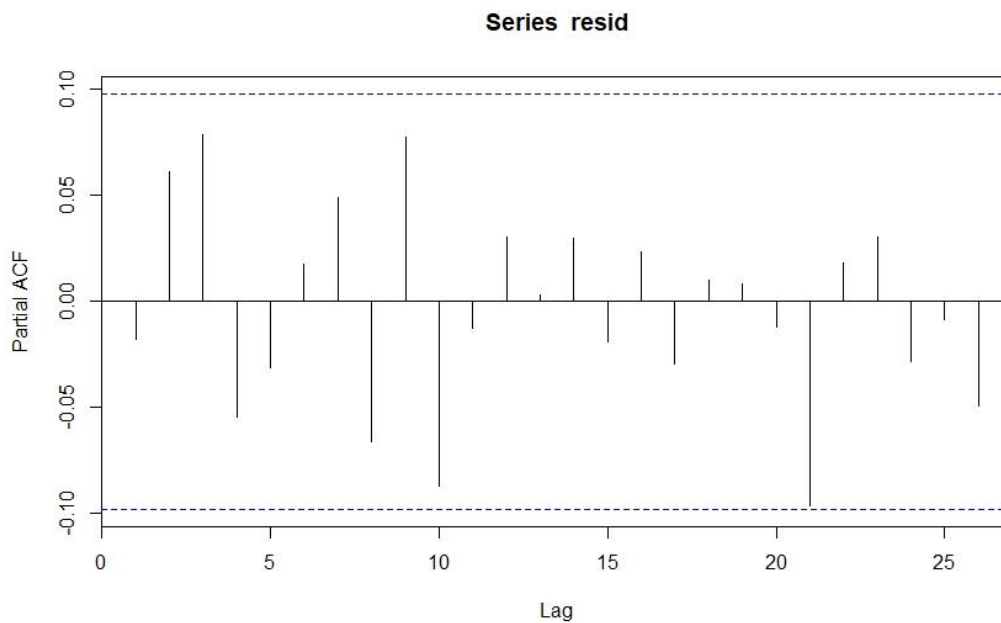
Abychom se přesvědčili o správnosti, pro názornost si ukážeme autokorelační funkce pro rezidua modelu.

Graf 6 - ACF reziduí odhadnutého modelu



Zdroj: vlastní zpracování – výstup programu RStudio

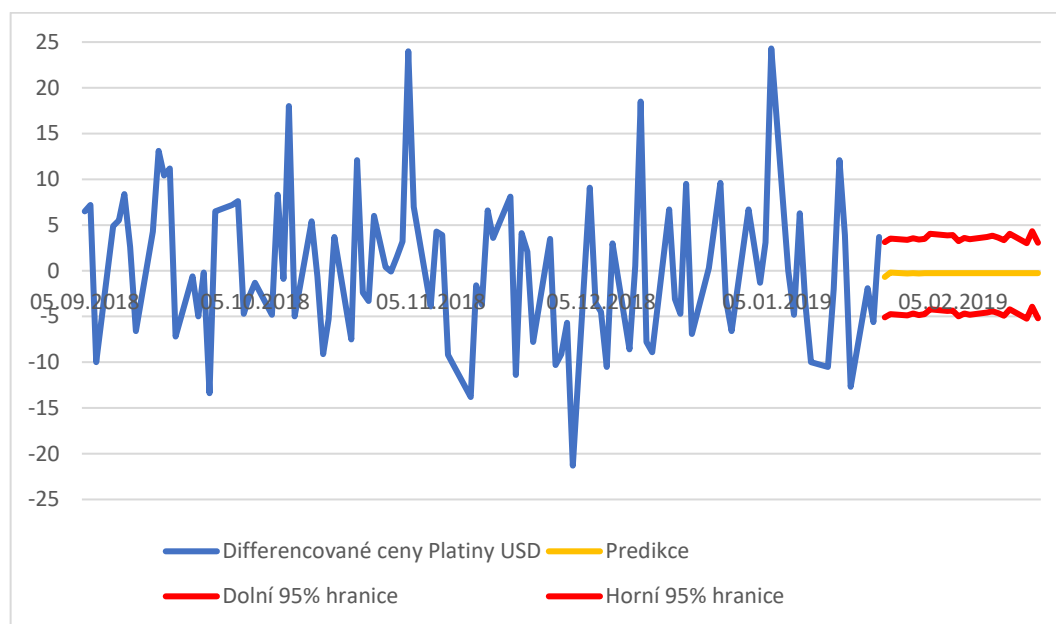
Graf 7 - PACF reziduí odhadnutého modelu



Zdroj: vlastní zpracování – výstup programu RStudio

Jak je z obrázků vidět, znázorněné autokorelační funkce nevyvracejí to, že byl model odhadnut chybně. Můžeme tedy přistoupit k samotným predikcím modelu, které získáme pomocí funkce `ugarchboot()`. Predikovat budeme pro názornost 20 budoucích hodnot. Pro přehlednost je v grafu zobrazeno jen 100 historických pozorování.

Graf 8 - Grafické znázornění predikce diferencovaných hodnot



Zdroj: vlastní zpracování

Z grafu se může zdát, že mají předpovědi konstantní charakter, nicméně diferencovaná data oscilují kolem nuly, nevykazují trend, a tak se dá očekávat, že budou dál pokračovat stejným směrem. Vypočítané intervaly spolehlivosti se zdají být poněkud malé, nicméně s rostoucím počtem pozorování klesá hodnota směrodatné odchylky, tudíž se zmenšuje i interval spolehlivosti.

Směrodatná odchylka byla spočítána na základě toho, že u každé predikce známe hodnotu směrodatné odchylky σ . Interval spolehlivosti se spočítá *predikční hodnota* $\pm 2 \cdot \sigma$

V následující tabulce je možné se podívat na tabulku konkrétních předpovědí.

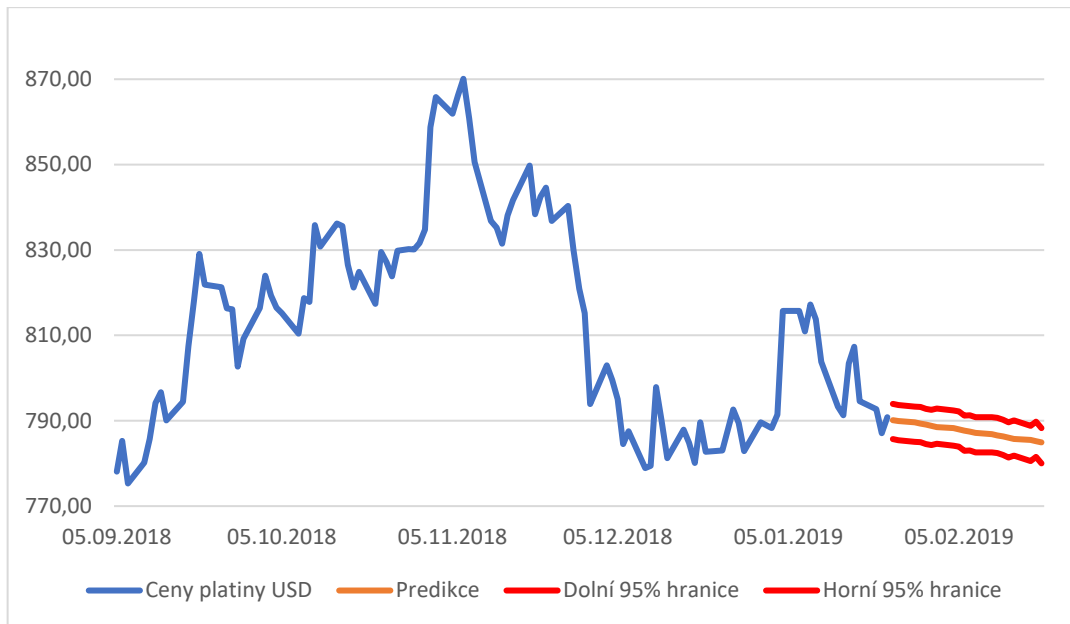
Tabulka 5 - Tabulka predikčních hodnot s 95 % intervalem spolehlivosti

Datum	Predikce	Dolní 95% hranice	Horní 95% hranice
24.01.2019	-0,673978	-5,098971	3,149213
25.01.2019	-0,1954636	-4,73058	3,517912
28.01.2019	-0,2940373	-4,864959	3,383841
29.01.2019	-0,2737312	-4,681212	3,567894
30.01.2019	-0,2779142	-4,833623	3,415791
31.01.2019	-0,2770525	-4,752418	3,497302
01.02.2019	-0,27723	-4,204806	4,045219
04.02.2019	-0,2771935	-4,388046	3,862285
05.02.2019	-0,277201	-4,358953	3,891683
06.02.2019	-0,2771994	-4,993817	3,257123
07.02.2019	-0,2771998	-4,675348	3,575897
08.02.2019	-0,2771997	-4,810462	3,441086
11.02.2019	-0,2771997	-4,552166	3,699686
12.02.2019	-0,2771997	-4,428165	3,82399
13.02.2019	-0,2771997	-4,634835	3,617622
14.02.2019	-0,2771997	-4,907254	3,345506
15.02.2019	-0,2771997	-4,223275	4,029786
18.02.2019	-0,2771997	-5,234842	3,018521
19.02.2019	-0,2771997	-3,928314	4,325349
20.02.2019	-0,2771997	-5,194727	3,059237

Zdroj: vlastní zpracování

Více nám prozradí, když aplikujeme predikční hodnoty na původní časovou řadu. Nejprve je nutné časovou řadu převést zpět do původní podoby, což je přesně opačný postup než při diferencování časové řady.

Graf 9 - Znázornění predikce na původních datech



Zdroj: vlastní zpracování

Z tohoto grafu je zřejmé, že se jedná o sestupný trend. Vytváření takovéto predikce je možné automatizovat a vytvořit tak indikátor pro predikce budoucích cen vybraných instrumentů. O tom, zda je možné takovéto predikování použít pro obchodování vybraných komodit se budeme zabývat v následující kapitole pojednávající o backtestu strategie.

4.2 Obchodní systém

Obchodní systém by měl pracovat následovně. Každý obchodní den vytvořit predikce na aktuální den a dle výsledků se buď rozhodnout k obchodu (nákup nebo prodej), či nikoliv. Využívat by měl čistě námi vytvořený indikátor ARIMA/GARCH. Před tím, než vůbec začneme s testováním na reálných datech, je třeba zjistit, zda by takový systém vůbec mohl být použitelný. Přesně to by nám měl prozradit backtest na námi vybraných komoditách.

4.2.1 Backtest

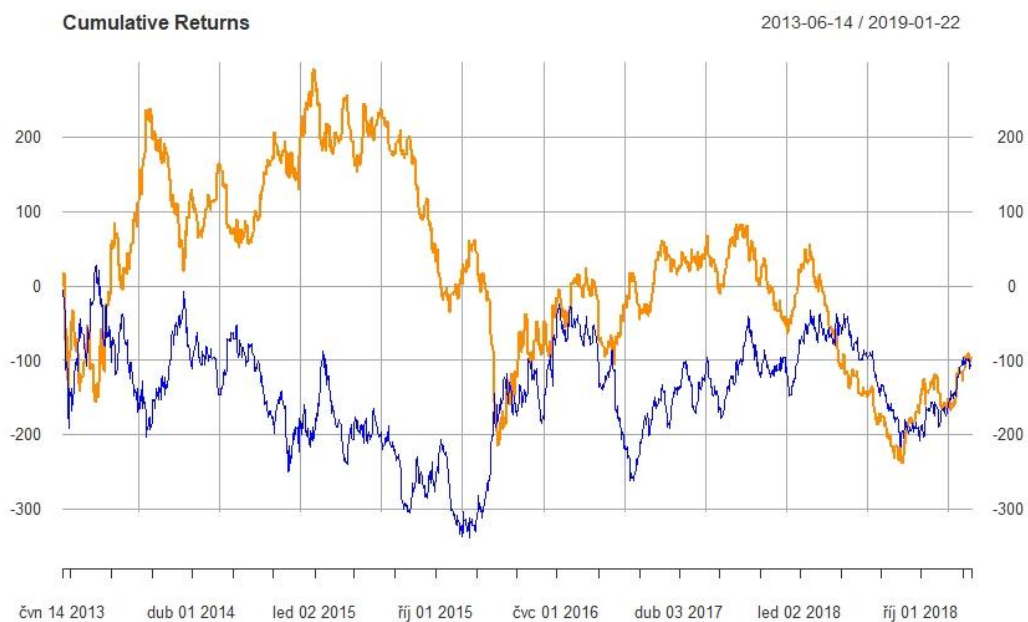
Backtest slouží, jak už název napovídá, k testování obchodního systému na historických datech. Platí zde základní předpoklad, že pakliže byl systém úspěšný v minulosti, mohl by být úspěšný i v současnosti.

Backtesting umožňuje v testeru strategií samotná platforma MetaTrader4, která do backtestu započítává všechny aspekty trhu. Tedy výši fixního spreadu i pohyby trhu, které se odehrály v mezičase. Pro vysvětlení ráno se otevrou ceny většinou na předchozí zavírací ceně, pak se pohybují oběma směry, a nakonec skončí na zabírací ceně pro dané období. Může se tak stát, že za jedno období (pro nás den) dojde k výraznému pohybu dolů, cena se ale v průběhu dne dorovná, a nakonec skončí nad otevírací cenou. Tento pohyb může strategii značně ovlivnit. Bohužel testování v přímo v platformě MetaTrader4 není možné použít, jelikož platforma v testovacím režimu nekomunikuje s „okolním světem“, tedy ani s R ne. Proto budeme pro test používat pouze zavírací ceny vybraných komodit, a vytvářet ho budeme přímo v prostředí R. Zohledněny nebudou ani spready, ani intradenní pohyby trhu.

Cílem celého backtestování je porovnání strategie „Buy and hold“ s námi vytvořenou strategií. Buy and hold je nejjednodušší investiční strategie, kdy se na začátku instrument koupí a prodá se až na konci. Náš obchodní systém každý den vždy nakoupí a prodá přesně tak, jak predikuje náš prediktivní ARIMA/GARCH indikátor. Pakliže je obchodní systém úspěšnější než pouhá investiční strategie, je z dlouhodobého hlediska vhodná pro obchodování. Denní historické záznamy byly k dispozici přibližně od října 2011. Backtest byl prováděn na celém období, které bylo k dispozici. Každý test měl za úkol provést přibližně 1500 predikcí ARIMA/GARCH modelu a každý trval necelé tři hodiny.

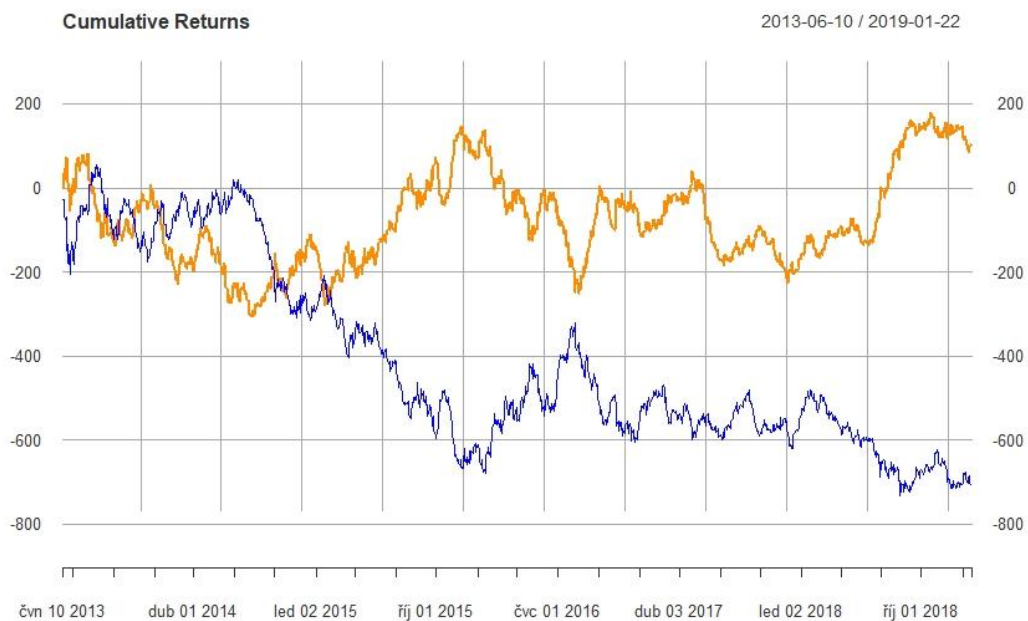
Následující obrázky zobrazují úspěšnost naší strategie v porovnání s Buy and hold strategií.

Graf 10 - Porovnání B&H strategie s ARIMA/GARCH strategií na zlatu



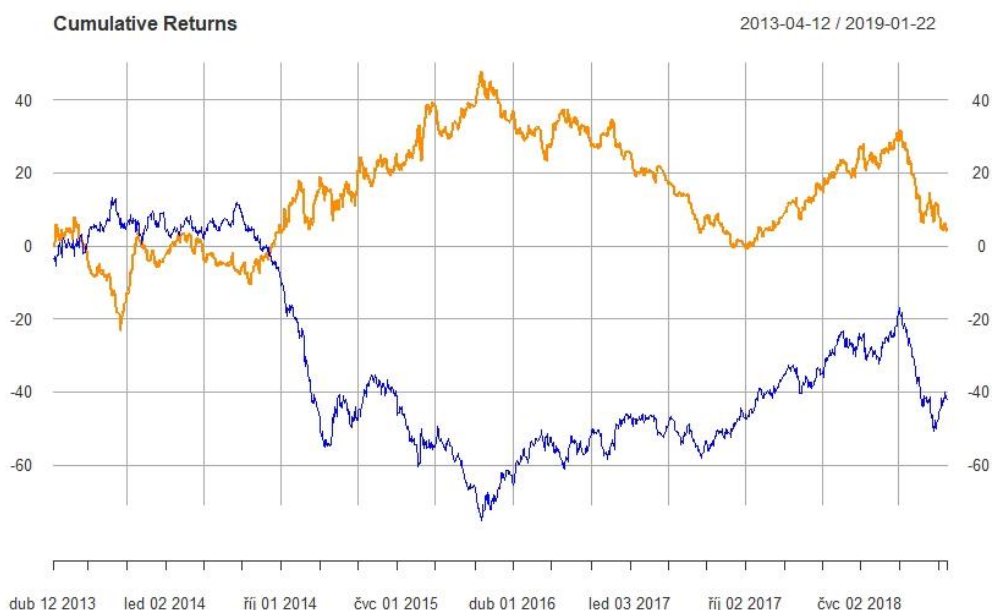
Zdroj: vlastní zpracování – Výstup programu RStudio

Graf 11 - Porovnání B&H strategie s ARIMA/GARCH strategií na platině



Zdroj: vlastní zpracování – Výstup programu RStudio

Graf 12 - Porovnání B&H strategie s ARIMA/GARCH strategií na ropě



Zdroj: vlastní zpracování – Výstup programu RStudio

Na grafech porovnávají strategie, kde modrá křivka znázorňuje kumulovaný zisk strategie Buy and hold a oranžová křivka znázorňuje kumulované příjmy ARIMA/GARCH strategie.

Až na zlato vychází naše strategie na první pohled lépe. Z dlouhodobého hlediska tedy můžeme považovat strategii za ziskovou. Za povšimnutí stojí to, že ačkoliv skoro na každé komoditě, pro kterou strategii testujeme, existuje dlouhé období růstu a také dlouhé období poklesu. Může to znamenat to, že, například strategie ARIMA/GARCH byla na ropě od roku 2016 rok a půl ztrátová. Přesto je její celkový výsledek kladný. Pouze zlato ze všech backtestů zůstalo víceméně na stejné hodnotě jako strategie Buy and Hold. Výhodou je, že minimální hodnota kumulovaných zisků na naší strategii je vyšší, než minimum na strategii B&H. To znamená, že maximální ztráta naší strategie nebyla tak vysoká jako u strategie B&H. Maximální ztráta je krom zisku jeden z nejdůležitějších parametrů při vyhodnocování úspěšnosti strategie.

Celkově tedy můžeme říci, že takto vytvořená strategie skutečně může fungovat. V následující kapitole se budeme zabývat funkčností systému na reálných datech.

4.2.2 Jazyk MQL a způsob použití

Abychom mohli testovat systém na reálných datech, musíme mít platformu, která umí stáhnout data, zadat obchodní příkaz a zároveň máme možnost doprogramovat si do ní něco svého. Tyto požadavky splňuje platforma MetaTrade4 a broker, který nám zprostředkovává data a obchodní příkazy X-Trade broker.

Jelikož je jádro systému naprogramované v jazyce R, je třeba použít nástroj, který umí propojit obě použité technologie. Bylo použité rozhraní mt4R od 7bit, které je ke stažení na stránce <https://www.forexfactory.com/showthread.php?t=260422>. Toto rozhraní umožňuje spustit jednu i více relací R a přenášet data tam i zpět mezi platformou MetaTrader4 a R. Relace R jsou v tomto případě spuštěny na pozadí jako samostatné procesy. Nespouští se tedy celý program R. Komunikace probíhá přes synchronní volání API. Potřebné soubory pro instalaci tohoto rozhraní jsou mt4R.dll a mt4R.mqh. Tyto soubory se nainstalují dle instrukcí do souborového adresáře platformy MetaTrader4. Dll knihovna se stará o samotnou výměnu dat mezi dvěma různými programy a soubor mqh poskytuje samotné rozhraní pro zadávání příkazů skrz platformu MetaTrader4. [32] Veškeré funkce pro komunikaci s R jsou velmi přehledně a stručně popsány v samotném souboru mqh.

Způsob vytvoření indikátoru spočívá v pár krocích. Nejprve je třeba provést základní nastavení indikátoru. Poté se spustí funkce `init()`, spustí se relace R a inicializují se tzv. buffery. To jsou datové pole, díky kterým je vykreslován vlastní indikátor. Použit se také mohou z důvodu předávání nevykreslovaných dat expertnímu algoritmu, který dále s indikátorem pracuje. Dále se dokola spouští funkce `start()`, dokud nedojde k ukončení činnosti indikátoru vnějším zásahem. Pro ukončení činnosti se volá funkce `deinit()`, která uzavře všechny relace, aby nezatěžovali paměť počítače. [33]

Pro tvorbu indikátoru je klíčová právě funkce `start()`, kde probíhá celá logika výpočtu.

4.2.3 Aplikace na reálných datech

Obchodní systém byl podroben testu na reálných datech na demoúčtu, který poskytuje společnost X-Trade Brokers. Tohoto brokera jsme vybrali díky jeho popularitě a možnosti obchodování komodit, což ne každý broker poskytuje. Zároveň poskytuje podporu pro

platformu MetaTrader4. Testování započalo 23.9.2019 až do data 18.10.2019. Systém byl následující. Každý den ráno provést pomocí námi vytvořeného indikátoru analýzu vývoje pro aktuální den a na tomto základě provést akci. To znamená provést nákup či prodej, nebo se rozhodnout pro daný den danou komoditu neobchodovat. Nejpozději každý večer byly všechny otevřené pozice uzavřeny. Co se týče řízení pozic, nastavován byl pouze příkaz Stop loss jako ochrana před nečekaným prudkým pohybem proti obchodní pozici. Zároveň, jelikož se při backtestu ukázalo, že délka referenčního období má také významný vliv na úspěšnost strategie, při každé analýze byla délka referenčního období upravena v intervalu 300 – 500 tak, aby byla autokorelace reziduí co nemější.

Velikosti pozic byly zvoleny tak, aby každý jednotlivý bod měl stejnou hodnotu u všech zvolených komodit. Jeden bod by se měl rovnat jednomu dolaru. To platí u všech s výjimkou platiny. Je to dáno výpočtem velikosti pozice. [13]

$$\text{Velikost pozice se počítá v letech dle vzorce: } Loty = \frac{1}{\text{Hodnota bodu pro daný instrument}}$$

Graf 13 - Hodnota 1 bodu na 1 lot pozice

Instrument	Hodnota 1 bodu na 1 lot
Gold	1
Oil	10
Platinum	15

Zdroj: X-Trade Brokers – Specifikační tabulka obchodních instrumentů

Směrodatnější je informace o bodech pohybu, tedy kolik bodů daná strategie zvládla, či nezvládla vydělat. Jeden bod je nejmenší možný pohyb ceny. Výhodou hodnocení strategie dle bodů je směrodatnější v tom ohledu, že je jedno s jakým kapitálem a jakou pákou se obchoduje. Bod se vždy přizpůsobí nastavené míře rizika, kterou si obchodník zvolí. [34]

Graf 14 - Minimální bodové pohyby na vybraných komoditách

Instrument	Bod (minimální pohyb ceny)
Gold	0,01
Oil	0,01
Platinum	0,1

Zdroj: X-Trade Brokers – Specifikační tabulka instrumentů

Při testování na reálných datech musíme zahrnout i velikost spreadů jednotlivých instrumentů. Spread je rozdíl mezi nákupní a prodejní cenou. V případě našeho účtu byly spready fixní a jejich velikost si můžeme prohlédnout v následující tabulce. [34]

Graf 15 - Velikosti spreadu u vybraných komodit

Instrument	Velikost spreadu v bodech
Gold	35
Oil	4
Platinum	74

Zdroj: Obchodní platforma MetaTrader4 obchodní účet X-Trade Brokers

Spread de facto udává, kolik bodů musí obchod minimálně vydělat, aby byl alespoň na nule.

Veškerý zdrojový kód vytvořeného indikátoru i zdrojový kód programu R včetně popisujících komentářů jsou uvedeny v příloze.

4.2.4 Pravidla obchodního systému

Každý obchodní systém musí mít pevně nastavená pravidla, dle kterých se obchoduje. Z hlediska řízení pozic a managementu bez pravidel nemůže fungovat žádný systém. Námí vytvořený systém budeme obchodovat na demo účtu s 100 000 fiktivními USD počátečním kapitálem. Systém a pravidla dle kterých bude fungovat jsou následující:

- Při otevření pozice se nastaví příkaz stop loss na nejbližší hranici opakovaného odrazu trhu, tzv. S/R zónu.

- Velikost pozice se uzavře tak, aby jeden bod pohybu obchodovaného instrumentu nestál více než 1 USD.
- Pozice se otevře jen tehdy, když bude p-hodnota Ljung-box testu větší než 0.10, tedy 2x větší než hladina významnosti.
- Pozice se otevře jen tehdy, když bude predikce pohybu ceny alespoň +/-0.12% z aktuální ceny.
- Ráno mezi 7 – 8 hodinou proběhnou analýzy na aktuální den a na základě analýz se obchody otevřou či neotevřou.
- Večer mezi 20 – 21 hodinou se všechny otevřené pozice uzavřou, pokud nebyly ukončeny dříve příkazem stop loss.

Obchodní systém bude z největší části obchodovat dle technické analýzy. Fundamentální analýzu nevyužije a z psychologické analýzy využívá jen psychologické hranice pohybu S/R zóny pro umístění příkazu stop loss. Z hlediska money managementu se nastavili tak nízké hodnoty pozic, že i kdyby celý měsíc obchodních dní bylo ztrátových, což by se čistě teoreticky mohlo stát, stále by bylo s čím obchodovat. Cíl takto nastaveného money managementu je mít tak velkou rezervu, aby se i přes extrémní neúspěch nemuseli ukončit testy předčasně.

5. Výsledky obchodování na reálných datech

Z měsíčního testu na reálných hodnotách byl veden obchodní deník a vytvořen report přímo z platformy MetaTrader4. Kompletní report z obchodování je uveden v příloze včetně obchodního deníku.

Výsledky obchodního systému můžeme dál analyzovat a zjišťovat, kde jsou mezery systému, které se pak můžeme snažit vylepšit. Pro analýzu výsledků můžeme použít program, který je navržen pro analyzování obchodních účtů, Quant Analyzer 4, nebo můžeme provádět analýzu na základě obchodního deníku. Obchodní deník si vytváří každý sám a je doporučováno si jej vytvářet, neboť platforma samotná není schopna vyreportovat, na jakém základě byl, či nebyl onen obchod uzavřen. Z obchodního deníku tedy získáme další doplňující údaje pro analýzu.

5.1 Analýza výsledků strategie

Za daný měsíc jsme pomocí zvolené strategie uzavřeli 38 obchodů. Celkově zůstala v plusu 205 bodů, což není moc. Na následujícím screenu z programu Quant Analyzer 4 vidíme podrobné výsledky strategie.

Obrázek 1 - Screen z programu Quant Analyzer 4 podrobné analýzy výsledků obchodování

TOTAL PROFIT	# OF TRADES	SHARPE RATIO	PROFIT FACTOR	RETURN / DD RATIO	WINNING PERCENTAGE
\$ 80.93	38	0	1.02	0.03	44.74 %
PROFIT IN PIPS	DRAWDOWN	% DRAWDOWN	DAILY AVG PROFIT	MONTHLY AVG PROFIT	AVERAGE TRADE
82410 PIPS	\$ 3133.18	3.11 %	\$ 4.05	\$ 80.93	\$ 236.14
YEARLY AVG PROFIT	ANNUAL % / MAX DD %	R EXPECTANCY	R EXPECTANCY SCORE	STR QUALITY NUMBER	SQN SCORE
\$ 971.16	0.03	0.01 R	4.59 R	0.03	0.05
YEARLY AVG % RETURN					
0.96 %					
CAGR					
0.08 %					

STATS					
Strategy					
Wins / Losses Ratio	0.81	Payout Ratio (Avg Win/Loss)	1.26	Average # of Bars in Trade	0
AHPR	0	Z-Score	-3.09	Z-Probability	99.9 %
Expectancy	2.13	Deviation	\$ 498.94	Exposure	0 %
Stagnation in Days	22	Stagnation in %	84.62 %		

Trades					
Gross Profit	\$ 4527.22	# of Wins	17	# of Losses	21
Largest Win	\$ 1606	Gross Loss	\$ -4446.29	Average Win	\$ 266.31
Avg Consec Wins	2.83	Largest Loss	\$ -1160	Max Consec Wins	11
		Avg Consec Loss	4.2	Max Consec Losses	6
				Avg # of Bars in Wins	0
				Avg # of Bars in Losses	0

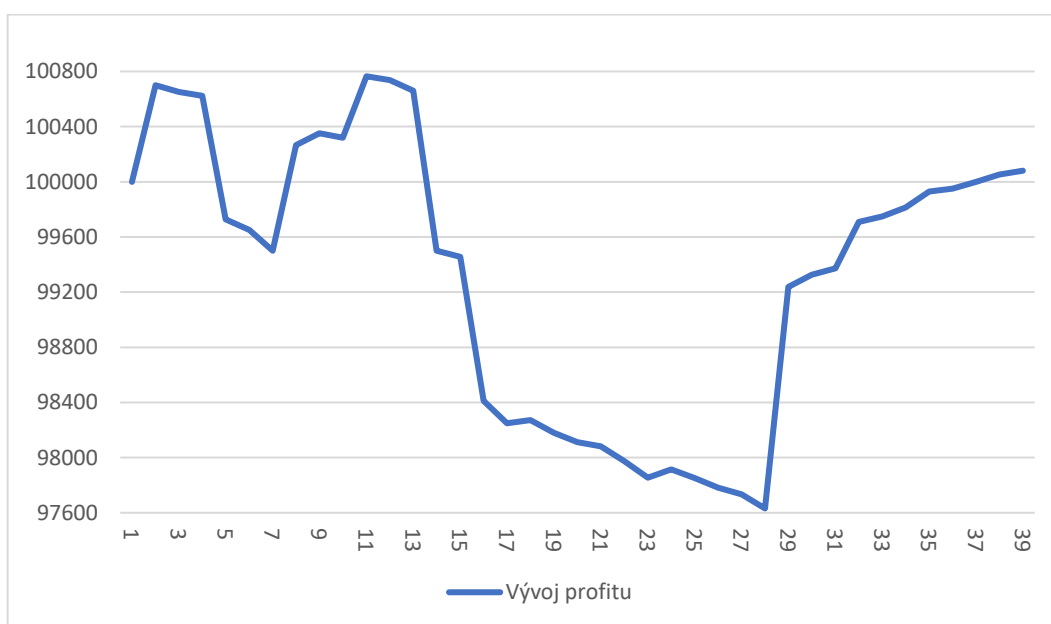
MONTHLY PERFORMANCE (\$)													
Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	YTD
2019	0	0	0	0	0	0	0	0	-499.61	580.54	0	0	80.93

Zdroj: vlastní zpracování – report programu Quant Analyzer 4

Můžeme si všimnout, že ačkoliv celková strategie byla celkově 80.9 USD v profitu, její maximální ztráta (Drawdown) za testovací období činila 3133.1 USD, kterou později dorovnala. Krom profitu a maximální poklesu je pro nás důležité i procento úspěšnosti obchodů, které zde je 44.74 % ziskových obchodů.

Pro názornost si ukážeme graf vývoje profitu naší strategie.

Graf 16 - Vývoj kapitálu při testování strategie na reálných datech



Zdroj: vlastní zpracování

S touto strategií bychom co se týče výsledků moc spokojeni nebyli. Má velké výkyvy. Maximální propad je násobně vyšší než zisk. To může být dáno malým úsekem testování. Jak jsme si ukazovali v backtestu, zisková strategie může být ztrátová i několik měsíců v řadě, takže soudit dle jednoho měsíce není až tak směrodatné. Nicméně pojďme si výsledky strategie rozebrat více do hloubky v analýze jednotlivých komodit.

5.2 Analýza jednotlivých komodit

Každá komodita je něčím zvláštní a její chování je vždy trochu odlišné od chování ostatních komodit. Rozebereme si proto výsledky každé komodity zvlášť, abychom si mohli udělat lepší obraz o fungování naší obchodní strategie.

5.2.1 Zlato

Na obrázku číslo 2 vidíme podrobný přehled výsledků strategie pouze pro zlato.

Obrázek 2 - Screen z programu Quant Analyzer 4 podrobné analýzy výsledků obchodování zlata

TOTAL PROFIT		# OF TRADES	SHARPE RATIO	PROFIT FACTOR	RETURN / DD RATIO	WINNING PERCENTAGE					
\$ 868.01		9	0.11	1.28	0.39	66.67 %					
PROFIT IN PIPS	86800 PIPS	DRAWDOWN	\$ 2204	% DRAWDOWN	2.18 %	DAILY AVG PROFIT	\$ 72.33	MONTHLY AVG PROFIT	\$ 868.01	AVERAGE TRADE	\$ 785.33
YEARLY AVG PROFIT	\$ 10416.12	ANNUAL % / MAX DD %	0.4	R EXPECTANCY	0.09 R	R EXPECTANCY SCORE	10.08 R	STR QUALITY NUMBER	0.33	SQN SCORE	0.03
YEARLY AVG % RETURN	10.44 %										
CAGR	0.87 %										

STATS					
Strategy					
Wins / Losses Ratio	2	Payout Ratio (Avg Win/Loss)	0.64	Average # of Bars in Trade	0
AHPR	0.1	Z-Score	-0.41	Z-Probability	65.91 %
Expectancy	96.45	Deviation	\$ 1125.63	Exposure	0 %
Stagnation in Days	21	Stagnation in %	84 %		

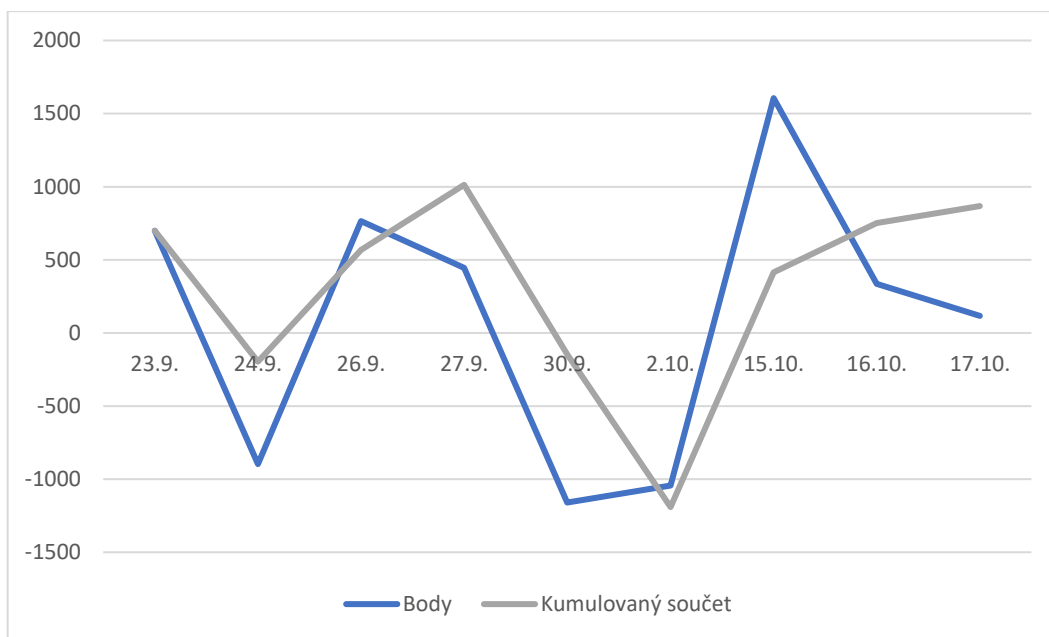
Trades					
# of Wins	6	# of Losses	3	# of Cancelled/Expired	0
Gross Profit	\$ 3968.01	Gross Loss	\$ -3100	Average Win	\$ 661.34
Largest Win	\$ 1606	Largest Loss	\$ -1160	Max Consec Wins	3
Avg Consec Wins	2	Avg Consec Loss	1.5	Max Consec Losses	2
		Avg # of Bars in Wins	0	Avg # of Bars in Losses	0

MONTHLY PERFORMANCE (\$)													
Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	YTD
2019	0	0	0	0	0	0	0	0	-146.99	1015	0	0	868.01

Zdroj: vlastní zpracování – report programu Quant Analyzer 4

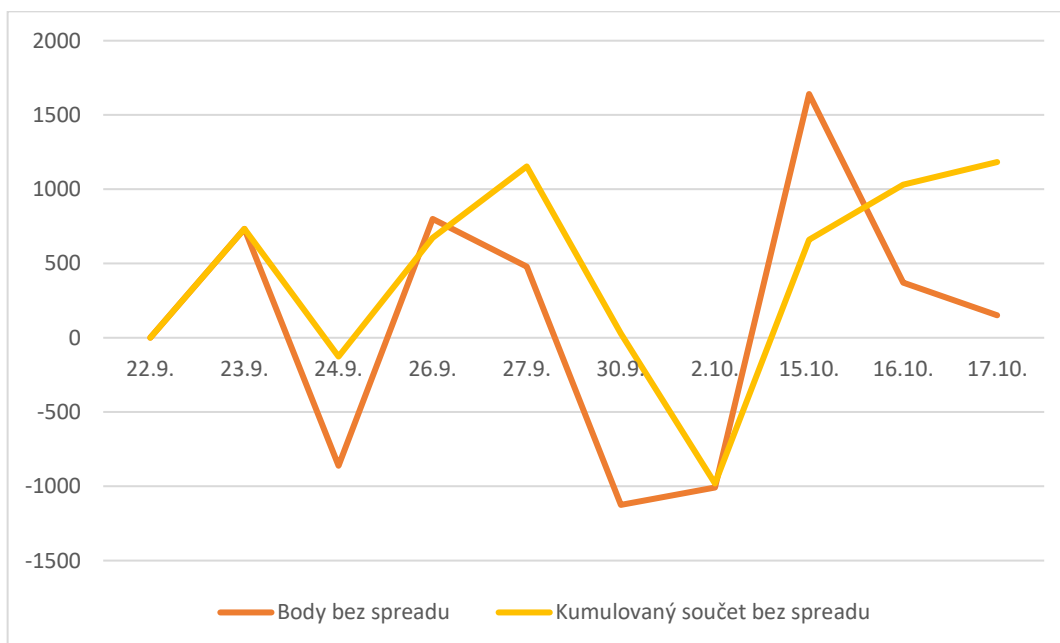
Obchodní systém na zlatu vykazuje poměrně vysokou míru úspěšnosti obchodů 66.67 %. Přitom maximální pokles byl 2204 USD, což není úplně ideální, ale celkově obchody na zlatu vydělaly 868 USD při celkovém počtu devíti obchodů. Mohlo by být zajímavé zakreslit si do grafu vliv spreadů, které při backtestu nebyly brány v úvahu.

Graf 17 - Bodový výsledek obchodování zlata na reálných datech



Zdroj: vlastní zpracování

Graf 18 - Bodový výsledek obchodování zlata na reálných datech s potlačením spreadu



Zdroj: vlastní zpracování

Na grafech je vidět, že spread měl na výsledek zlata také vliv. Konkrétně 315 USD by připadlo jako spreadový poplatek brokerovi za umožnění obchodu. Nicméně průběh obou grafů je téměř totožný, takže můžeme konstatovat, že spread na zlatě nijak neohrožuje funkčnost našeho systému.

5.2.2 Ropa

Na ropě bylo uskutečněno celkem 12 obchodů. Pojďme se podívat na detailní výsledky z výstupu QuantAnalyzer4.

Obrázek 3 - Screen z programu Quant Analyzer 4 podrobné analýzy výsledků obchodování ropy

TOTAL PROFIT	\$ 49.02	# OF TRADES	12	SHARPE RATIO	0	PROFIT FACTOR	1.15	RETURN / DD RATIO	0.25	WINNING PERCENTAGE	58.33 %
PROFIT IN PIPS	4900 PIPS	DRAWDOWN	\$ 192.98	% DRAWDOWN	0.19 %	DAILY AVG PROFIT	\$ 3.27	MONTHLY AVG PROFIT	\$ 49.02	AVERAGE TRADE	\$ 57.92
YEARLY AVG PROFIT	\$ 588.24	ANNUAL % / MAX DD %	0.26	R EXPECTANCY	0.06 R	R EXPECTANCY SCORE	9.11 R	STR QUALITY NUMBER	0.22	SQN SCORE	0.04
YEARLY AVG % RETURN	0.6 %										
CAGR	0.05 %										

STATS

Strategy					
Wins / Losses Ratio	1.4	Payout Ratio (Avg Win/Loss)	0.82	Average # of Bars in Trade	0
AHPR	0	Z-Score	0.42	Z-Probability	33.72 %
Expectancy	4.09	Deviation	\$ 83.36	Exposure	0 %
Stagnation in Days	23	Stagnation in %	88.46 %		

Trades					
Gross Profit	\$ 372.01	# of Wins	7	# of Losses	5
Largest Win	\$ 88	Gross Loss	\$ -322.99	Average Win	\$ 53.14
Avg Consec Wins	1.75	Largest Loss	\$ -107.99	Max Consec Wins	4
		Avg Consec Loss	1.25	Max Consec Losses	2
				Avg # of Bars in Wins	0
				Avg # of Bars in Losses	0

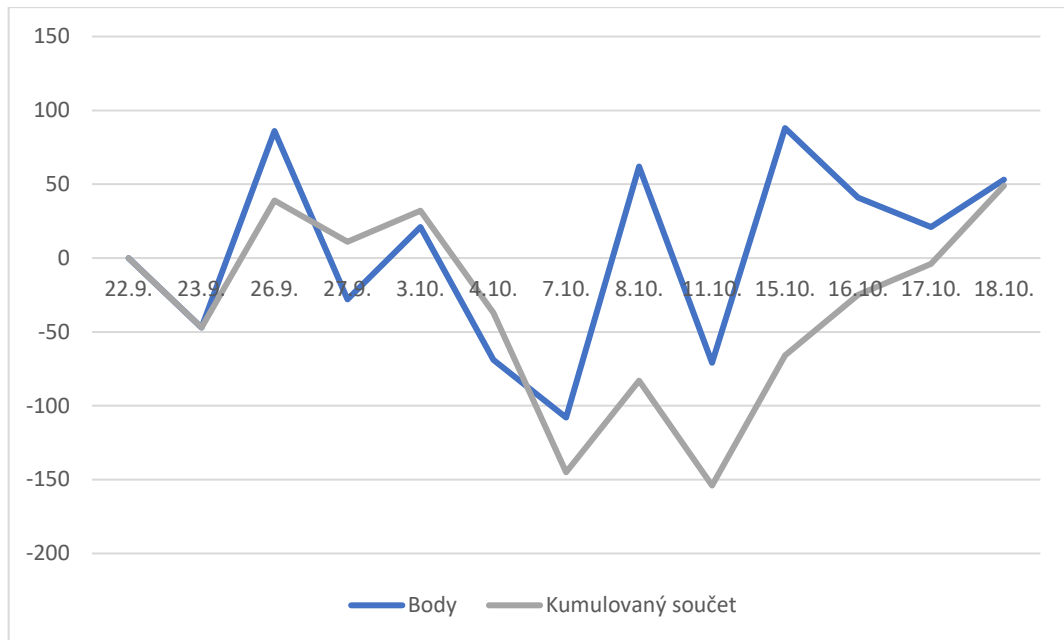
MONTHLY PERFORMANCE (\$)

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	YTD
2019	0	0	0	0	0	0	0	0	11	38.02	0	0	49.02

Zdroj: vlastní zpracování – report programu Quant Analyzer 4

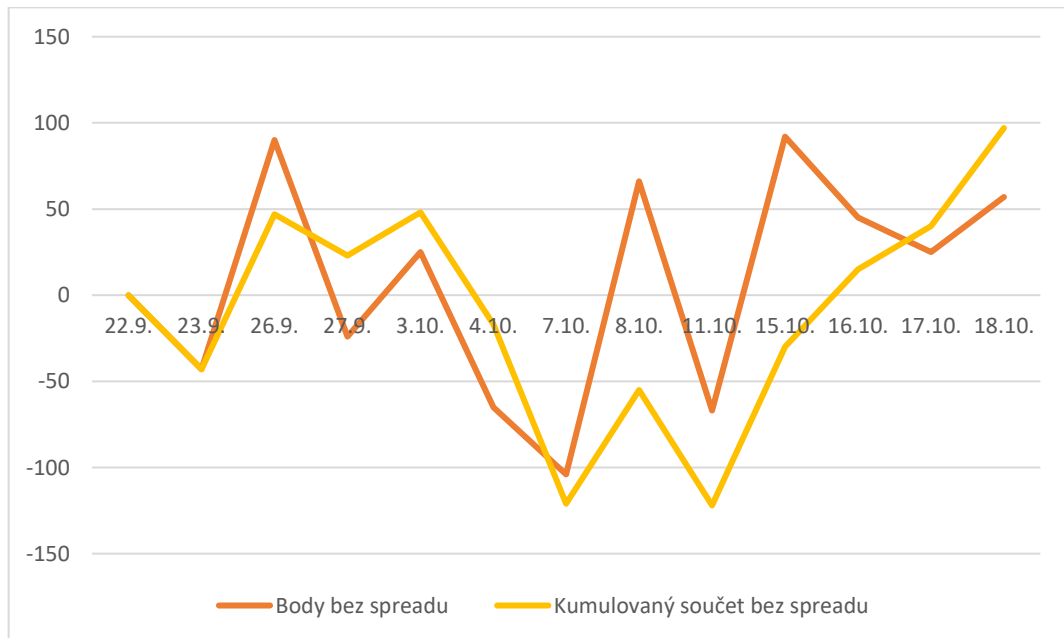
Maximální propad je zde znatelně menší a to 193 USD. Je to dáno volatilitou ropy. Ropa byla mnohem méně volatilní než zlato, či platina. Naproti tomu celkový zisk je pouze 49 USD. Procento úspěšných obchodů je zde 58.33 %, což je pozitivní. Opět se podíváme na grafy vývoje a porovnáním se systémem, kde by nebyly žádné spready.

Graf 19 - Bodový výsledek obchodování ropy na reálných datech



Zdroj: vlastní zpracování

Graf 20 - Bodový výsledek obchodování ropy na reálných datech s potlačením spreadu



Zdroj: vlastní zpracování

Ropa měla spread jen 4 body, proto i oba grafy mají téměř stejný průběh. Opět jako u zlata se dá konstatovat, že spready mají na náš obchodní systém založený na ropě minimální vliv.

5.2.3 Platina

Na platině bylo provedeno nejvíce obchodů s celkovým počtem 17 obchodů. Pojďme se podívat na její detailní statistiku.

Obrázek 4 - Screen z programu Quant Analyzer 4 podrobné analýzy výsledků obchodování platiny

TOTAL PROFIT \$ -836.1	# OF TRADES 17	SHARPE RATIO -0.76	PROFIT FACTOR 0.18	RETURN / DD RATIO -0.82	WINNING PERCENTAGE 23.53 %
PROFIT IN PIPS -929000 PIPS	DRAWDOWN \$ 1023.3	% DRAWDOWN 1.02 %	DAILY AVG PROFIT \$ -44.01	MONTHLY AVG PROFIT \$ -836.1	AVERAGE TRADE \$ 71.21
YEARLY AVG PROFIT \$ -10033.2	ANNUAL % / MAX DD % -0.82	R EXPECTANCY -0.62 R	R EXPECTANCY SCORE -127.45 R	STR QUALITY NUMBER -3.1	SQN SCORE -1.07
YEARLY AVG % RETURN -10.08 %					
CAGR -0.84 %					

STATS					
Strategy					
Wins / Losses Ratio	0.31	Payout Ratio (Avg Win/Loss)	0.59	Average # of Bars in Trade	0
AHPR	-0.05	Z-Score	-4.02	Z-Probability	100 %
Expectancy	-49.18	Deviation	\$ 137.06	Exposure	0 %
Stagnation in Days	26	Stagnation in %	100 %		

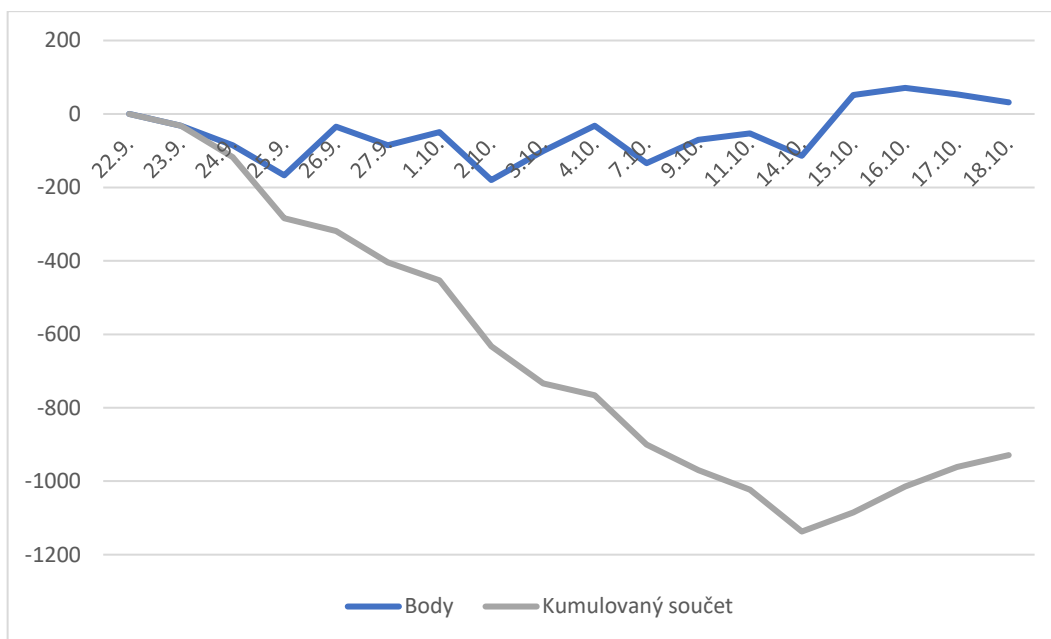
Trades					
# of Wins	4	# of Losses	13	# of Cancelled/Expired	0
Gross Profit	\$ 187.2	Gross Loss	\$ -1023.3	Average Win	\$ 46.8
Largest Win	\$ 63.9	Largest Loss	\$ -162	Max Consec Wins	4
Avg Consec Wins	4	Avg Consec Loss	13	Max Consec Losses	13
		Avg # of Bars in Wins	0	Avg # of Bars in Losses	0

MONTHLY PERFORMANCE (\$)													
Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	YTD
2019	0	0	0	0	0	0	0	0	-363.62	-472.48	0	0	-836.1

Zdroj: vlastní zpracování – report programu Quant Analyzer 4

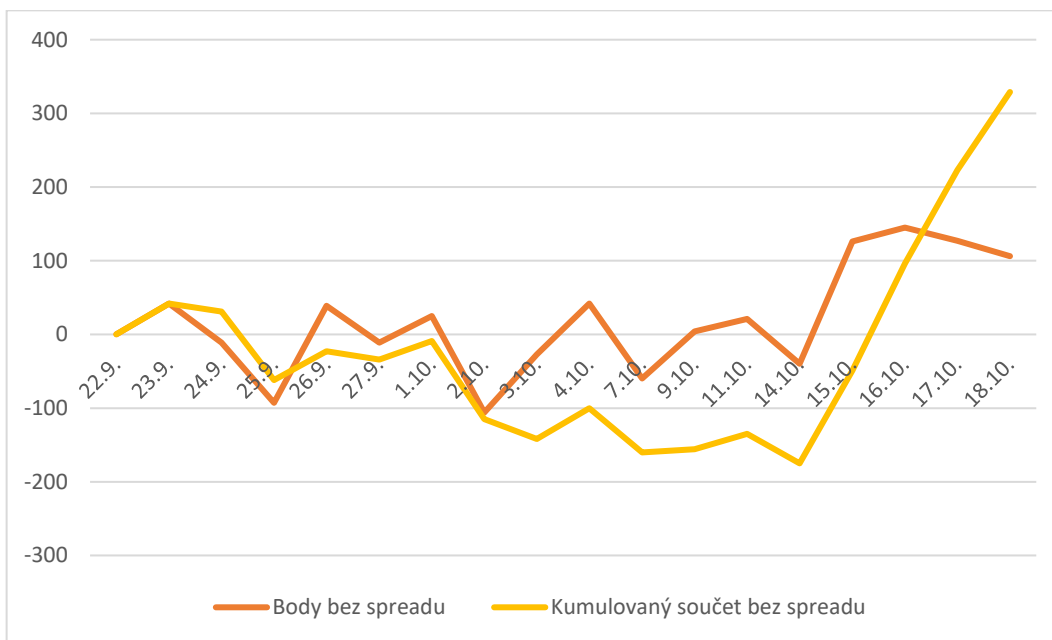
Platina skončila v „červených“ číslech, avšak vykazuje nejmenší propad ze všech testovaných komodit. Maximální propad u platiny je 1023 USD, což je uspokojivé. Co je neuspokojivé je procento úspěšnosti dané strategie a samozřejmě profit. Vysvětlit by to mohly následující grafy.

Graf 21 - Bodový výsledek obchodování platiny na reálných datech



Zdroj: vlastní zpracování

Graf 22 - Bodový výsledek obchodování platiny na reálných datech s potlačením spreadu



Zdroj: vlastní zpracování

Platina má u námi používaného brokera poměrně velký spread 74 bodů. Jak je vidět na grafech. Zde tento spread značně ovlivňuje výsledek strategie. Platina z tohoto důvodů pro takovýto typ obchodování není vhodná. Spread by musel být znatelně nižší, což vzhledem k tomu, že platina není tak hodně obchodovaná jako ostatní uvedené komodity dost možná není zcela možné. Obecně totiž platí, že čím více se daná komodita obchoduje, tím menší spread je. Broker si tím kompenzuje provozní náklady. Další možností, jak uplatnit strategii ARIMA/GARCH je na delším časovém období, například týdenní grafy. Zde je ale velmi obtížné testování strategie, jelikož by trvalo příliš dlouho.

5.3 Analýza obchodního deníku

Obrovská výhoda obchodního deníku je v tom, že jsme do něj zaznamenávaly jen údaje o provedených, ale také o neprovedených obchodech a všechny údaje, které nám ARIMA/GARCH indikátor spočítal. Zejména údaje o parametrech, zvoleném referenčním období, p-hodnotu Ljung-boxova testu a hodnotu predikce. Díky tomuto obchodnímu deníku můžeme zjistit nejen to, jak si obchodní systém vedl, ale i to, kde jsou jeho slabiny a jak ho případně zlepšit.

Pomocí kontingenčních tabulek zjistíme, které prvky mají vliv na úspěšnost pozic. V tomto testu nebudeme zohledňovat výši bodů, jelikož vykazují velký rozptyl, ale pokusíme se vyfiltrovat ty potenciálně ziskové od těch ztrátových. Prvním krokem je tedy převést výsledky do jednotlivých pozic na dummy proměnou, kde -1 znamená ztrátu, 0 znamená, že se obchod neuskutečnil a 1 znamená, že daná pozice skončila v zisku.

Jako první se budeme zabývat, jestli mělo referenční období vliv na to, zda pozice skončila v zisku či nikoliv.

Tabulka 6 - Kontingenční tabulka počtu úspěšných, neúspěšných a neuskutečněných obchodů vzhledem k referenčnímu období

Počet z Výsledek	Popisky sloupců			
Popisky řádků	-1	0	1	Celkový součet
320			1	1
340	3		2	5
360	2	3	1	6
380	3	4	5	12
400	3	5	1	9
420	3	2	1	6
440	4	4	1	9
460		1	1	2
480	1		2	3
500	2		2	4
Celkový součet	21	19	17	57

Zdroj: vlastní zpracování

Na vertikální ose jsou délky referenčních období a na horizontální ose jsou výsledky daných obchodů. Hodnoty ukazují počty výskytu události. Můžeme vidět, že nepoužívanější období bylo 380 záznamů a zároveň vypadá jako nejlepší možná volba, jelikož má nejlepší poměr mezi ztrátovými a ziskovými obchody. Počet pozorování je ale nízký na to, abychom mohli s vysokou pravděpodobností určit, jestli existuje referenční období, které je správnější než ostatní.

Zajímavá situace je u parametrů. Vytvořili jsme kontingenční tabulku s vertikální osou p a horizontální osou q , jejichž hodnoty jsou součty výsledků strategie.

Tabulka 7 - Kontingenční tabulka vlivu odhadnutých parametrů ARMA na úspěšnost modelu.

Součet z Výsledek	Popisky sloupců						
Popisky řádků	0	1	2	3	4	5	Celkový součet
0		3					3
1	2	2	1				5
2			-4	-3	-1		-8
3	0		1	-2			-1
4					-2	0	-2
5					-2	1	-1
Celkový součet	2	5	-2	-5	-5	1	-4

Zdroj: vlastní zpracování

Strategie s mnohem větší pravděpodobností vybere takové parametry, které jsou blízké hlavní diagonále např. ARMA(1,1), ARMA(2,2), atd. Jen zřídka se stane, že se od sebe oba parametry významně liší. V našem případě se maximálně lišili o 3 body v konkrétním případě ARMA(3, 0). Tato informace by mohla být užitečná při optimalizaci, kde by nemusel probíhat celý dvojitý cyklus na zjištění optimálních parametrů, ale stačily by použít hodnoty sousedící v určitém maximálním rozpětí. Dle obchodního deníku jsou ziskovější výpočty ARMA modelů nižších řádů, což by potvrzovalo tvrzení Arlt, J. a Arltová, M. [1], že model ARMA(1,1) je v praxi nejčastěji užívaný. Dávalo by tedy smysl neobchodovat modely s parametry vyšších řádů.

Jako poslední parametr, který stojí za to, podívat se na něho blíže je p-hodnota Ljung-box-testu, která nám testuje přítomnost autokorelace reziduí. V obchodním systému se sice obchod neuzavíral, pokud tato hodnota vycházela nižší než 0.1, což je 2krát více, než je obvyklá hladina významnosti, ale i tak z obchodního deníku je patrné, že jsou nejspíš úspěšnější obchody, který mají hladinu významnosti nastavenou ještě o dost výš. Nejprve jsme všechny p-hodnoty zaokrouhlili na jedno desetinné místo a vynásobili deseti, abychom je mohli rozřadit do kategorií a sumarizovat kontingenční tabulkou.

Tabulka 8 - Kontingenční tabulka vlivu p-hodnoty Ljung-Box testu rozdělených do deseti kategorií na úspěšnost obchodů

Počet z Výsledek	Popisky sloupců			
Popisky řádků	-1	0	1	Celkový součet
1	2	5	1	8
2	4		3	7
3	3	3	1	7
4	3	1		4
5	1	2	2	5
6	1			1
7	1	2	3	6
8	4		1	5
9	1	2		3
10	1	4	6	11
Celkový součet	21	19	17	57

Zdroj: vlastní zpracování

Nejvíce kladných výsledků vykazuje desátá kategorie, tedy p-hodnota s hodnotou 9.5 a více. Čistě teoreticky bychom se mohli zaměřit jen na ni a ostatní obchody neobchodovat. 57 obchodů na 10 kategorií je málo na to, aby takovýto závěr mohl mít větší váhu. Proto jsme p-hodnoty rozdělili pouze na dvě skupiny. První skupina obsahuje kategorii 1-5 a druhá skupina obsahuje kategorie 6-10. Zajímá nás tedy, zda posunout hladinu významnosti až na 0.5 z původních 0.1.

Tabulka 9 - Kontingenční tabulka vlivu p-hodnoty Ljung-Box testu rozdělených do dvou kategorií na úspěšnost obchodů

Počet z Výsledek	Popisky sloupců			
Popisky řádků	-1	0	1	Celkový součet
0	12	9	5	26
1	9	10	12	31
Celkový součet	21	19	17	57

Zdroj: vlastní zpracování

Úspěšnost menších p-hodnot než 0.54 je téměř přesně opačná než úspěšnost vyšších p-hodnot. U menších p-hodnot bylo ziskových pouze 5 ze 17 provedených obchodů, což je 29,4% pravděpodobnost, že pokud uzavřeme obchod s nižší p-hodnotou než 0.54, obchod zůstane v zisku. U vyšších byla situace úspěšnosti 12 ziskových obchodů z 21 provedených. Tedy 57.1% pravděpodobnost, že pokud se rozhodneme obchodovat vyšší p-hodnotu, tak skončí v zisku. Na tomto základě bych pro budoucí testy doporučoval posunout p-hodnotu Ljung.Boxova testu na minimální hodnotu 0.54, jinak obchod neuzavírat.

6. Závěr

Na základě metodických postupů tvorby ARIMA/GARCH modelu se nám podařilo vytvořit algoritmus v R, který je schopný vytvářet tyto modely a predikce pouze na základě dodaných dat.

Vytvořený algoritmus byl dále upraven, abychom byli schopni provést backtest. Ten byl prováděn na všech dostupných historických datech stažených z platformy a ukázal, že strategie obchodování pomocí ARIMA/GARCH modelu může být z dlouhodobého pohledu zisková.

To nám umožnilo tento algoritmus propojit s obchodní platformou MetaTrader4 pomocí API a vytvořit vlastní indikátor, který automaticky vyhodnocuje reálná data. Na základě tohoto indikátoru byl vytvořen obchodní systém, kterým jsme se zabývali.

Obchodní systém jako takový by z dlouhodobého hlediska sice mohl fungovat. Svědčí o tom provedené backtesty v R, avšak v krátkodobém časovém horizontu úspěšnost systému test na reálných datech v obchodní platformě nepotvrdil.

Hlavním přínosem práce pro veřejnost je uvedení do problematiky obchodních systémů fungujících na základě pokročilých prediktivních modelů. Svě si zde najde každý, kdo se zajímá o obchodování pomocí matematicko-statistických modelů, případně jen obecně o obchodování na burze. Dalším přínosem je vysvětlení fungování a praktické využití ARIMA/GARCH modelu. V neposlední řadě je přínosem i to, že na takovém systému je možné dále pracovat a rozvíjet ho.

Díky této práci jsme si osvojili funkcionalitu lineárních a nelineárních procesů, pronikly do základů programování v R i MQL4 a prakticky práce vedla k lepšímu pochopení fungování prediktivních modelů na tak volatilním a nevyzpytatelném trhu, jako jsou burzovní data komodit.

6.1 Doporučení pro zlepšení

Pro kvalitnější otestování, před samotným obchodováním s reálnými penězi, by systém musel fungovat mnohem déle. Testem jsme si však ověřili, že systém je funkční a použitelný. Zároveň s sebou přinesl další poznatky, které je možné využít pro vylepšení našeho obchodního systému.

- Pokud bychom např. vylučovali všechny obchody, pro které vychází p -hodnota Ljung-Box testu menší než 0.54, úspěšnost našeho systému by se mohla značně zvýšit.
- Na základě testu na reálných datech bychom mohlo optimalizovat i rychlost algoritmu, který je momentálně velmi výpočetně náročný. Docílili bychom toho úpravou cyklů pro hledání správných parametrů, jelikož jsme zjistili, že řády parametrů p a q se od sebe většinou velmi neliší.

Z provedených backtestů je zřejmé, že strategie může v průběhu svého fungování i rok a půl ztrácet a teprve poté se stane ziskovou. Má dlouhé ziskové i ztrátové období. Doporučení pro další vývoj spočívá v trendech strategie. Odfiltrování negativních trendů výkonnosti strategie by mohlo přinést velice zajímavé výsledky obchodního systému.

7. Seznam použitých zdrojů

- [1] ARLT, Josef a Markéta ARLTOVÁ. *Ekonomické časové řady: [vlastnosti, metody modelování, příklady a aplikace]*. 1. Praha: Grada, 2007. ISBN 978-80-247-1319-9.
- [2] MOLNÁR, Karol. *Moderní síťové technologie*. Brno, 2008. Dostupné také z: <http://www.utko.feec.vutbr.cz/~molnar/mmos/fronty.pdf>. Skripta. FEKT VUT Brno.
- [3] KŘIVÝ, Ivan. *Analýza časových řad*. Ostrava, 2006. Dostupné také z: http://www.informatika-osu.czechian.net/files/is/ancas/ANCAS_DiV.pdf. Skripta. Ostravská univerzita.
- [4] CIPRA, Tomáš. *Finanční ekonometrie*. 2., upr. vyd. Praha: Ekopress, 2013. ISBN 978-80-86929-93-4.
- [5] PERELLI, Roberto. *Introduction to ARCH & GARCH models*. Illinois, 2001. Dostupné také z: <http://www.econ.uiuc.edu/~econ472/ARCH.pdf>. Skripta. University of Illinois.
- [6] CIPRA, Tomáš. *Analýza časových řad s aplikacemi v ekonomii: celostátní vysokoškolská učebnice pro stud. matem.-fyz. fakult studijních oborů 11 Fyzikálně matematické vědy*. 1. Praha: Státní nakladatelství technické literatury, 1986. ISBN 99-00-00157-X.
- [7] JUSKO, Martin. *Modely volatility ARCH a GARCH*. Praha, 2009. Dostupné také z: <http://invenio.nusl.cz/record/283012>. Bakalářská práce. Univerzita Karlova. Vedoucí práce Hudecová Šárka.
- [8] WILSON, Michael. *A Historical View of Network Traffic Models*. St. Louis, 2006. Dostupné také z: https://www.cse.wustl.edu/~jain/cse567-06/ftp/traffic_models2.pdf
- [9] YANG, Yuhong. *Can the Strengths of AIC and BIC Be Shared? A Conflict between Model Identification and Regression Estimation* [online]. Oxford University Press: Biometrika, 2005, 2005(92) [cit. 2019-05-26]. ISSN 1464-3510. Dostupné z: <https://www.jstor.org/stable/20441246?seq=1>
- [10] Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models. *Journal of the American Statistical Association*, 1970, 65(332). Dostupné také z: <https://www.jstor.org/stable/2284333?seq=1>
- [11] *On a Measure of Lack of Fit in Time Series Models* [online]. Oxford University Press: Biometrika, 1978, 65(2) [cit. 2019-05-20]. ISSN 1464-3510. Dostupné z: <https://www.jstor.org/stable/2335207?seq=1>

- [12] NESNÍDAL, Tomáš a Petr PODHAJSKÝ. *Obchodování na komoditních trzích: průvodce spekulanta. 2., rozš. vyd.* Praha: Grada, 2006. Investice. ISBN 80-247-1851-0.
- [13] HARTMAN, Ondřej. *Začínáme na burze: jak uspět při obchodování na finančních trzích - akcie, komodity a forex.* Brno: BizBooks, 2013. ISBN 978-80-265-0033-9.
- [14] BP Statistical Review of World Energy 2018. In: *BP* [online]. USA: BP, 2018 [cit. 2018-11-19]. Dostupné z: <https://www.bp.com/content/dam/bp/en/corporate/pdf/energy-economics/statistical-review/bp-stats-review-2018-oil.pdf>
- [15] DROZEN, František. *Drahé kovy a kameny.* Praha: Vysoká škola ekonomická, 2000. Edice základů oceňování majetku. ISBN 80-245-0078-7.
- [16] Major countries in global mine production of platinum from 2012 to 2017 (in metric tons). In: *Statista* [online]. Německo: Statista, 2018 [cit. 2018-11-19]. Dostupné z: <https://www.statista.com/statistics/273645/global-mine-production-of-platinum/>
- [17] PLATINUM QUARTERLY: Q2 2018. World Platinum Investment Council. In: *Oxford: World Platinum Investment* [online]. Oxford: SFA, 2018 [cit. 2019-11-19]. Dostupné z: https://www.platinuminvestment.com/files/943838/WPIC_Platinum_Quarterly_Q2_2018.pdf
- [18] ROGERS, Jim. *Žhavé komodity: jak může kdokoliv investovat se ziskem na světových trzích.* Praha: Grada, 2008. Investice. ISBN 978-80-247-2342-6.
- [19] JÍLEK, Josef. *Finanční a komoditní deriváty v praxi. 2., upr. vyd.* Praha: Grada, 2010. Finanční trhy a instituce. ISBN 978-80-247-3696-9.
- [20] WGCa. How much gold has been mined? Word Gold Council [online]. Velká Británie: Word Gold Council, 2018 [cit. 2018-11-19]. Dostupné z: <https://www.gold.org/about-gold/gold-supply/gold-mining/how-much-gold-has-been-mined>
- [21] WGCb. Supply and demand. *Word Gold Council* [online]. Velká Británie: Word Gold Council, 2018 [cit. 2018-11-19]. Dostupné z: <https://www.gold.org/data/gold-supply-and-demand/gold-market-chart>
- [22] Annual Report 2017. *Barrick* [online]. Toronto: Word Gold Council, 2018 [cit. 2019-11-26]. Dostupné z: <https://barrick.q4cdn.com/788666289/files/annual-report/Barrick-Annual-Report-2017.pdf>

- [23] HARTMAN, Ondřej. *Jak se stát forexovým obchodníkem: naučte se vydělávat na měnových trzích*. Praha: FXstreet, c2009. ISBN 978-80-904418-0-4.
- [24] REJNUŠ, Oldřich. *Teorie a praxe obchodování s cennými papíry*. Praha: Computer Press, 2001. Business books (Computer Press). ISBN 80-722-6571-7.
- [25] ELDER, Alexander. *Tradingem k bohatství: psychologie, obchodní systémy, money management*. Tetčice: Impossible, c2006. Knihovna úspěšného obchodníka. ISBN 80-239-7048-8.
- [26] REJNUŠ, Oldřich. *Finanční trhy*. 4., aktualiz. a rozš. vyd. Praha: Grada, 2014. Partners. ISBN 978-80-247-3671-6.
- [27] PERSON, John L. *A complete guide to technical trading tactics*. Hoboken, N.J., c2004. ISBN 04-715-8455-X.
- [28] DVOŘÁK, Roman. *Trading strategie: moderní styl obchodování na burze: včetně popisu třech funkčních trading strategií*. Brno: Computer Press, 2008. ISBN 978-80-251-2240-2.
- [29] JANÁČ, Radek. *Automatické obchodní systémy, aneb, Forex Robot*. 1. Brno: Tribun EU, 2016. ISBN 978-80-263-1055-6.
- [30] ŠTÝBR, David, Petr KLEPETKO a Pavlína ONDRÁČKOVÁ. *Začínáme investovat a obchodovat na kapitálových trzích*. Praha: Grada, 2011. Finance pro každého. ISBN 978-80-247-3648-8.
- [31] *The R Project for Statistical Computing* [online]. 2019 [cit. 2019-11-27]. Dostupné z: <https://www.r-project.org/>
- [32] 7BIT. MQL4 -> R-Project - Interface Library. In: *Forex Factory* [online]. 2010 [cit. 2019-11-27]. Dostupné z: <https://www.forexfactory.com/showthread.php?t=260422>
- [33] *MQL5* [online]. Limassol: MQL5, 2019 [cit. 2019-11-27]. Dostupné z: <https://www.mql5.com/>
- [34] *X-Trade Brokers* [online]. Varšava: X-Trade Brokers, 2019 [cit. 2019-11-27]. Dostupné z: <https://www.xtb.com/cz>

8. Přílohy

Příloha A	Zdrojový kód programu R
Příloha B	Zdrojový kód indikátoru ARIMA/GARCH na platformě MT4
Příloha C	Report obchodování na demo účtu z platformy MT4
Příloha D	Obchodní deník