EKONOMENTRIJSKI MODELI ZA PREDVIĐANJE RAZINE CIJENA I VOLATILNOSTI NAFTE

Perasović, Vicko

Master's thesis / Diplomski rad

2016

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Split, Faculty of economics Split / Sveučilište u Splitu, Ekonomski fakultet**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:124:372911

Rights / Prava: In copyright / Zaštićeno autorskim pravom.

Download date / Datum preuzimanja: **2024-03-29**

Repository / Repozitorij:

REFST - Repository of Economics faculty in Split

zir.nsk.hr

UNIVERSITY OF SPLIT

dabar

DIGITALNI AKADEMSKI ARHIVI I REPOZITORIJI

SVEUČILIŠTE U SPLITU
EKONOMSKI FAKULTET

DIPLOMSKI RAD

EKONOMENTRIJSKI MODELI ZA PREDVIĐANJE RAZINE CIJENA I VOLATILNOSTI NAFTE

Mentor:
doc. dr. sc. Arnerić Josip

Student:
Vicko Perasović

SADRŽAJ:

1. UVOD 4
   1.1. Problem istraživanja 4
   1.2. Predmet istraživanja 5
   1.3. Svrha i ciljevi istraživanja 7
   1.4. Metode istraživanja 7
   1.5. Doprinos istraživanja 8
2. KARAKTERISTIKE NAFTE 9
   2.1. Kratka povijest cijene, volatilnosti i geopolitičkih utjecaja nafte 9
   2.2. Vrste nafte 12
   2.3. WTI i Brent crude 14
   2.4. Futures ugovori nafte 15
   2.5. Odabir podataka 16
3. MODELIRANJE VREMENSKIH SERIJA 17
   3.1. Testiranje stacionarnosti i odabir razdoblja 17
   3.2. Sezonalnost 21
   3.3. Trend komponente 22
   3.4. AR model 24
   3.5. ARIMA model 26
   3.6. GARCH model 30
   3.7. Varijacije GARCH modela 34
4. ANALIZA FUNDAMENTALNIH FAKTORA 36
   4.1. Odabir fundalemntalnih faktora i njihovih vremenskih nizova 36
   4.2. Analiza korelacije 37
   4.2. Linearna regresija 39
5. ANALIZA I USPOREDBA REZULTATA 43
   5.1. Sažeti rezultati modela 43
6. ZAKLJUČAK 48

POPIS TABLICA **49**
POPIS SLIKA **49**
SAŽETAK **54**
SUMMARY **55**

1. UVOD

1.1. Problem istraživanja

Nafta ima nezamjenjivu vrijednost u svjetskoj ekonomiji. Osim tradicionalnih namjena nafte i njenih prerađevina, kao što je generiranje topline ili funkcije goriva u motornim vozilima, nafta se koristi i u druge, manje standardne svrhe. Akteri u svijetu financijskih tržišta naftu, kao i derivative izvedene iz nafte, koriste kao diversifikacijsko sredstvo svojih portfelja, predmet špekulacije za generiranje profita iznad razine zahtijevane stope povrata te u strategijama arbitraže.

Za razliku od poljoprivrednih proizvoda koji se mogu zamijeniti svake sezone ili metalnih proizvoda koji se mogu reciklirati gotovo u nedogled, fosilna goriva kao što je nafta su količinski ograničena.Strah od nestašice i strateška važnost nafte u globalnoj ekonomiji osiguravaju stabilan interes gotovo svih svjetskih vlada te poslovnih organizacija koje koriste naftu ili njene prerađevine kao inpute i outpute u svojim poslovnim modelima. Stoga performanse interesnih skupina znatno ovise o njihovim strategijama u svezi s naftom.

Neelastična priroda cijene nafte te visoka volatilnost u kratkom i dugom roku snažno utječu na globalna tržišta. Visoku volatilnost u usporedbi s drugim energetskim proizvodima istražio je Pindyck (1999). Posljedično, kroz povijest se javio znatan broj „naftnih šokova“ koji su imali dramatičan utjecaj na razvoj svjetske ekonomije. Dodatni dokaz o razvoju kompleksnih uporaba nafte se može pronaći na burzama financijskih derivativa. Burza na kojoj se trguje najvećim brojem Futures i Options ugovora nafte sa standardiziranim svojstvima svakog dana objavi statistike trgovanja ugovorima među kojima je i volumen trgovine. Od svih energetskih ugovora, WTI Crude, oblik nafte sa specifičnim karakteristika koja se transformira u SAD-u, ima najveći dnevni volumen trgovine, sa više od 200 000 ugovora dnevno za vrijeme pisanja ovog rada.

Karakteristike nafte i njena značajnost, koji su opisani u prethodnim odlomcima predstavljaju temeljnu motivaciju ovog istraživanja. Ako uzmemo za primjer proizvođače automobila ili zračne kompanije, njihova prodaja i troškovi značajno ovise o cijeni nafte. Mnoge poslovne organizacije neutraliziraju rizik naglih volatilnosti uzimanjem velikih pozicija u već spomenutim ugovorima. Druge kompanije špekuliraju koje će biti razine cijena u rasponu od nekoliko mjeseci do deset ili više godina. Ako ne mogu efikasno predvidjeti razinu cijena,

njihovo poslovanje može biti ugroženo. Posljednji nagli pad cijena 2014. godine, koji je gotovo prepolovio cijene nafte, je značajno utjecao na mnoge poslovne organizacije, ali i zemlje čija stabilnost ekonomije ovisi o izvozu ili uvozu nafte. Petrobras, Brazilska multinacionalna energetska korporacija, ujedno i najveća tvrtka u Južnoj Americi, morala je provesti značajne rezove da bih zadržala solventnost i dugoročnu stabilnost. Predviđanje cijene nafte je ključno s obzirom na visoku volatilnost u današnjoj poslovnoj okolini. Ekonometrijska analiza ima potencijal dati odgovore i smjernice za uspješno predviđanje.

1.2. Predmet istraživanja

Ovo istraživanje temelji se na određivanju i odabiru onih ekonometrijskih modela koji pokazuju najbolje rezultate u objašnjavanju buduće cijene i volatilnosti nafte. Postoji značajna literatura koja se bavi istraživanjem dinamike i volatilnosti nafte te se ujedno koristi raznovrsnim ekonometrijskim metodama. Alquist i Kilian (2007) u svom radu istražuju validnost jedne od temeljnih pretpostavki futures ugovora, a to je da su futures cijene nepristrani pokazatelji buduće trenutačne (spot) cijene. Lam (2013) istražio je i dinamiku nafte na temelju mnogih ekonometrijskih modela, uključujući ARIMA te GARCH i APARCH koji su korišteni za modeliranje reziduala.

Osim autoregresije i drugih metoda za modeliranje vremenskih serija, značajna je i literatura na području analize koristeći nezavisne varijable. Koristili su se regresijski i drugi modeli koji modeliraju cijenu nafte kroz varijable kao što su neto izvoz, stopa iskorištenosti rafinerija ili kamatne stope u SAD-u. Određivanje fundamentalnih faktora koji su kroz povijest imali značajan utjecaj na cijenu i volatilnost nafte možemo povući iz brojnih istraživačkih radova na ovom području. Jedan primjer je Kaufmannova (2004) primjena statističkih modela za procjenu veze između cijene nafte i različitih faktora razine proizvodnje, odnosno ponude nafte.

S druge strane, nafta ima značajan utjecaj kao nezavisna varijabla. Ferderer (1996) istražio je utjecaj cijene nafte na ekonomski rast te zaključio da je taj utjecaj negativne prirode. Park i Ratti (2008) istražili su asimetriju naftnih šokova na ekonomske aktivnosti. Iako navedeni pristup nije direktno vezan za predmet istraživanja ovog rada, mnoge poveznice za definiranje regresijskih nezavisnih varijabli mogu se pronaći u navedenim istraživanjima.

Volatilnost je dodatan aspekt koji će biti predmetom ovog istraživanja. Zašto je volatilnost važna? Visoke fluktuacije u cijenama nafte vrlo često imaju značajne utjecaje na opću ekonomiju. Primjerice, recesijska razdoblja nakon drugog svjetskog rata su prethođena visokim porastima cijena nafte u svijetu. Upravo to istražio je Hamilton (1983). Znanstvena literatura bazirana na volatilnosti nafte je bogata. Narayan i Narayan (2007) nastojali su modelirati volatilnost cijene nafte. U svom radu, Salisu i Fasanya (2012) pokazali su da asimetrični GARCH model pokazuje superiornije rezultate u objašnjavanju volatilnosti nafte što upućuje na dokaze velike financijske poluge na tržištu nafte. Još jedna interesantna varijacija na klasične ekonometrijske modele za procjenu volatilnost temelji se na futures ugovorima. Upravo to su istražili Fleming i Ostdiek (1999), ali nisu pronašli dovoljno čvrste dokaze koji vežu porast volatilnosti za porast volumena trgovanja i popularnosti ugovora.

Iako su se neke metode u prethodnim istraživanjima pokazale manje efikasnim od drugih, kao primjerice linearna regresija u radu Bashera i Sadorskya (2006), testirat ćemo intuitivno privlačne varijaciju regresije. Ovdje je fokus isključivo na predikcijske sposobnosti modela što daje naglasak na testiranje izvan uzorka, dok se u velikoj većini navedenih radova testiranje odvijalo unutar uzorka. Neki radovi su pokazali obećavajuće rezultate u predviđanju cijene nafte. Primjerice, Pyndick (1999) pokazao je predikcijsku moć odabranih modela za predviđanje nafte, ali se isti nisu pokazali dobri u slučaju drugih fosilnih goriva. Manescu i Robays (2014), u svom istraživanju za Europsku Centralnu Banku, pokazali su kvalitetu kombinacije četiri modela u predviđanju Brent-a: VAR, DSGE, futures ugovori i futures ugovori prilagođeni za rizik. Iznimno sveobuhvatno istraživanje na području predviđanja cijene nafte obavili su i Alquist, Kilian i Vigfusson (2011). Oni su na temelju empiričkih dokaza zaključili da nominalna i realna cijena nafte ne prate *random walk* proces te su detaljno istražili dvosmjerni utjecaj nafte i eksternih varijabli. Na temelju navedenih i drugih istraživanja izgradit ćemo naš pristup za definiranje potencijalnih modela predviđanja.

1.3. Svrha i ciljevi istraživanja

Cilj istraživanja je filtrirati one ekonometrijske metode koje najbolje objašnjavaju ponašanje i dinamiku cijene nafte te njenu volatilnost. Kako će se koristiti metode analize vremenskih nizova koje uključuju korištenje lagiranih vrijednosti, ali i komponente drugih vremenskih nizova kao što su tržišni indeksi i kamatne stope, istražit će se mnoge alternative predviđanja cijene nafte.

Kada se, prema definiranim kriterijima, odaberu najefikasniji pristupi i modeli, različiti sudionici naftnog tržišta mogu iskoristiti informaciju kao podlogu za donošenje kvalitetnih odluka. Istraživanje će utvrditi dinamiku ponašanja cijene i volatilnosti nafte te objasniti faktore koji imaju najznačajniji utjecaj i na taj način potencijalno opremiti donosioce odluka s ključnim znanjima o gotovo nezamjenjivom proizvodu u njihovom poslovanju.

1.4. Metode istraživanja

U teorijskom dijelu rada koristit će se opće znanstvene metode koje služe u znanstveno istraživačkim radovima kako bi se istražili rezultati znanstvenih istraživanja. Koristit će se: metoda indukcije, dedukcije, sinteze, dokazivanja i opovrgavanja, analize i komparativna metoda. Metoda indukcije: vrsta posrednog zaključka kod kojeg se polazi od pojedinačnog ka općem, odnosno ono što vrijedi za svaki pojedinačni slučaj jedne vrste vrijedi za cijelu vrstu. Metoda dedukcije (donošenje zaključaka od pojedinačnih sudova na opće). Metoda sinteze koja predstavlja jednu od osnovnih metoda spoznaje, odnosno spajanje jednostavnijih misli u složene. Metoda analize odnosno postupak istraživanja i objašnjavanja putem raščlanjivanja složenih zaključaka na znatno jednostavnije dijelove. Metoda dokazivanja i opovrgavanja je metoda koje se koristi kod dokazivanja točnosti postavljenih hipoteza. Komparativna metoda kojom će se pojedini pojmovi uspoređivati kako bi se utvrdile njihove sličnosti i razlike.

U empirijskom dijelu rada koristit će se statističke, matematičke i ekonometrijske metode s ciljem analize utjecaja odabranih faktora na razinu cijena i volatilnosti nafte korištenjem softverskog paketa R. Treba naglasiti i da će se u sklopu svakog analiziranog modela provjeravati njegova validnost testiranjem elemenata koji mogu imati utjecaj na iskrivljenost

modela kao što su heteroskedastičnost i autokorelacija koristeći se ekonometrijskim testovima kao što je Jarque Berra test.

1.5. Doprinos istraživanja

Već su spomenuti mnogi relevantni sudionici i korisnici naftnih proizvoda te njihovih financijskih derivativa. Također je napomenuta temeljna podjela aktera financijskih tržišta na arbitražere, špekulatore i hedgere. Bez obzira na orijentaciju, ciljeve i strategije, svaki od sudionika ima značajnu korist od poznavanja uzoraka ponašanja dvije najvažnije karakteristike nafte: cijene i volatilnosti. Uz to, predviđanje buduće volatilnosti i cijene može imati neizmjernu ekonomsku korist za pojedinog sudionika. Ovaj rad nastoji približiti temeljne karakteristike navedenih parametara nafte te pristup i metodologiju koja će biti bolja u odnosu na alternative u njihovom predviđanju. Empirijski će se istražiti i rangirati pristupi po odabranim kriterijima. Kada bih donosioci odluka u naftnoj industriji svoja predviđanja bazirali na kvalitetnoj ekonometrijskoj podlozi, bili bi mnogo spremniji na šokove i fluktuacije u cijeni nafte i u sklopu s tim, spremniji na najznačajnije rizike u svom poslovanju.

Dodatan potencijalan doprinos je orijentiran tržištu. Tržišta s velikim brojem sudionika koji na temelju kvalitetnih informacija donose precizne i informirane odluke, često su izvrsno procijenjena. To znači da cijene tih proizvoda točno reflektiraju njihovu intrizičnu vrijednost. U ovakvim tržištima potencijali za profitabilnost su mali ili nikakvi, ali zato su rizik i fluktuacije također iznimno niski. Kada bi se tržište nafte ponašalo na ovaj način, visoka volatilnost koja ima velike posljedice na svjetske ekonomije se ne bi ponavljala te bi ekonomski svijet bio mnogo stabilniji. Hamlinton (2009) je u svom radu na naftnim šokovima došao do zaključka da nedostaje sistematskog načina za predvidjeti recesiju te da bih se potencijal mogao nalaziti upravo u naftnim fluktuacijama. Možda upravo metodologija za predviđanje cijena nafte bazirana na ekonometrijskim performansama može doprinijeti stabilnosti financijskog sustava.

2. KARAKTERISTIKE NAFTE

2.1.Kratka povijest cijene, volatilnosti i geopolitičkih utjecaja nafte

Cijena nafte koja se prikazuje u suvremenim znanstvenim radovima, knjigama i vijestima, može se odnositi na dva fundamentalna proizvoda: spot cijena barela sirove nafte koja se koristi kao mjerilo (u sljedećem poglavlju su objašnjeni različiti oblici sirove nafte koji se koriste kao mjerila) te *futures* ugovori – ugovori koji za imatelja predstavljaju obvezu kupnje predodređenog iznosa barela nafte po unaprijed određenoj cijeni na točno definiran datum u budućnosti. Kako je danas volumen trgovanja futures ugovorima nafte među najvećima od svih futures ugovora te s obzirom da i dobavljači, kupci i drugi članovi opskrbnog lanca aktivno trguju ovim ugovorima u svrhu *hedge*-anja, špekuliranja i arbitraže, kroz ovaj rad ćemo pretežito promatrati cijenu nafte upravo kroz futures ugovore. Dodatna velika prednost u odnosu na spot cijenu je postojanje iznimno likvidnog tržišta gdje su precizno definirane kupovna i potražna cijena u svako vrijeme za različite mjesece isticanja ugovora. Prema tome možemo direktno iz tržišta vidjeti cijenu za ugovor nafte koji ističe u rujnu 2016. godine pa sve do posljednjeg koji ističe u prosincu 2024. godine.

Futures ugovori sirove nafte uvedeni su već u sedamdesetim godinama 20.stoljeća. Iranska Revolucija 1979. godine i Jomkipurski rat 1973. godine uzrokovali su ekstremne nesigurnosti u opskrbi nafte što je jako volatilno utjecalo na cijenu. Kako bi se tako visoke fluktuacije ublažile, *New York Mercantile Exchange (NYMEX)*burza uvela je prvi naftni futures proizvod pod nazivom WTI Crude Oil Futures. Cilj proizvoda bio je da služi kao sredstvo članovima naftne industrije koje mogu upotrijebiti za *hedganje* protiv neželjenih promjena u cijeni, odnosno osiguranje nesmetanog budućeg poslovanja.

Kroz znatan dio dvadesetog stoljeća, cijena sirove nafte bila je znatno regulirana kroz proizvodnu stranu i cjenovne kontrole. Sve do 2000. godine, kada je OPEC definirao cjenovni interval od $22 do $28, cijene su prelazile 30$ samo u slučaju ratnih konflikata u Bliskom istoku. Cjenovni interval regulacije OPEC je napustio 2005. godine.

Jedini vrlo dugoročni vremenski niz naftnih cijena, koji se proteže u razdbolju i prije drugog svjetskog rata, je prosjek izvorne spot cijene barela SAD-a prilagođenog za inflaciju. Na taj način možemo u daljnjem tekstu ove cjeline dobiti okvirne usporedive cijene na povijesnoj razini.

Prvi značajan geopolitički događaj koji je utjecao na cijenu nafte je već spomenuti Jomkipurski rat 1973. godine, napadanjem Izraela od strane Sirije i Egipta. Mnogi arapski dobavljači postavili su embargo kao reakciju na podršku SAD-a pružene Izraelu. Smanjenje proizvodnje za 5 milijuna barela dnevno od strane ovih zemalja, umanjenih za milion barela proizvodnje povećanih od strane drugih zemalja, predstavljalo je 7% svjetske proizvodnje nafte. Embargo je utjecao na porast cijene nafte za 400%.

Zatim, 1979. I 1980. godine, događanja u Iraku i Iranu dovela su do novih porasta cijene nafte. Iranska revolucija utjecala je na prosječno smanjenje između 2 i 2.5 milijuna barela dnevno. Post-revolucijsko razdoblje, uvjetovano invazijom Iraka 1980. godine, pospješilo je najveći porast cijene nafte u razdoblju poslije drugog svjetskog rata. Zbrojeni output barela Iraka i Irana u tom periodu iznosio je samo jedan milijun barela dnevno. Posljedično, svjetska proizvodnja barela pala je za 10%. Pad outputa pod utjecajem ova dva događaja doveo je do povećanja cijene nafte za 200%.

Porasti cijene u razdoblju od 1973. do 1981. godine bili bi i mnogo veći da nije bilo SAD-a, koji je uveo cjenovne kontrole u periodu nakon embarga. Rezultat je bio povećanje troška uvoza nafte za 50% u odnosu na domaću proizvodnju. To je efektivno dovelo do subvencioniranja potrošača u SAD-u od strane domaće naftne industrije, iz razloga što su imali manje prihode u odnosu na svjetske aktere naftne industrije. Efekti politike cjenovne kontrole su i dalje tema rasprave ali su izvan opsega ovog rada te ih nećemo razmatrati.

U razdoblju od 1980. do 1986. godine, output zemalja izvan OPEC kartela se povećao za 6 milijuna barela dnevno. Niža potražnja i viša ponuda uvjetovali su odluku OPEC-a da 1982. godine postave proizvodne kvote. Kvote su bile bezuspješne u stabilizaciji cijene nafte. Odluka Saudijske Arabije u kolovozu 1985. godine da vežu svoje cijene za spot cijenu nafte te poveća proizvodnju za tri milijuna barela u periodu kraćem od godinu dana dovelo je do naglog pada cijena nafte.

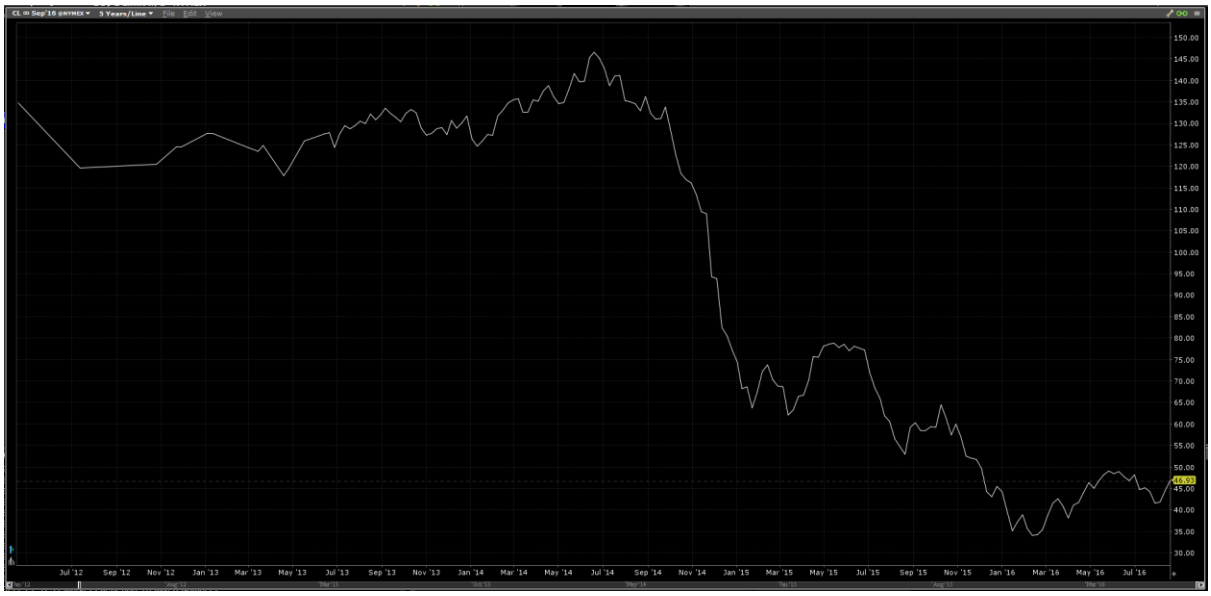
Sljedeći znatni geopolitički događaj dogodio se 1990. godine, kada je Irak pokreniju invaziju Kuvajta, započinjući Zaljevski rat. Do 1994. godine cijene nafte kontinuirano su padale te dotiču najnižu razinu 1973. godine. Tada su cijene promijenile smjer u kontinuirani rast na temelju povećanja svjetske proizvodnje nafte na 6.2 milijuna barela dnevno, slabljenja Ruske proizvodnje te snažnog jačanja ekonomije SAD-a. Sve do 1998. godine cijene nafte su rasle, kada su ekonomska kriza u Aziji u kombinaciji sa većim outputom OPEC-a pokrenili znatan pad u cijenama. Kao odgovor na to OPEC je smanjio kvote u dva navrata za 1.25 milijuna i

1.335 milijuna barela. To je dovelo do početka oporavka cijena u 1999. godini. Teroristički napad 11. rujna 2001. godine utjecao je na znatan pad cijena. Spot WTI cijene su pale za 35% do studenog. Za vrijeme svih navedenih događaja osim terorističkog napada, OPEC je regulirao cijene povećavanjem i smanjenjm kvota uz upitne rezultate.

2003. godine vojne intervencije u Iraku su započele. Gubitak proizvodnje u Iraku i Venezueli, kombinirani s povećanom proizvodnjom OPEC-a u svrhu zadovoljavanja svjetske potražnje, dovele su do smanjenja od preko četiri milijuna barela viška u odnosu na potražnju u samo jednoj godini. Kroz 2004. i 2005. godinu kapaciteti viška proizvodnje OPEC zemalja iznosile su samo milijun barela dnevno, što nije dovoljno za pokriti značajne prekide u outputu koje mogu uvjetovati geopolitički događaji. U tom periodu, svjetska potražnja iznosila je preko 80 milijuna barela dnevno. Dodatni utjecaji su bili slabi dolar te ubrzano povećanje Azijskih ekonomija. Sve navedeno dovelo je do znatnog porasta cijena do $50 po barelu nafte.

Godine 2008. je nastupila najduža recesija u SAD-u koja je ujedno zahvatila i cijeli svijet. Usljedio je znatan porast cijene nafte. Rekordnih $145.29 futures ugovora na NYMEX burzi postignuto je 3. srpnja 2008. godine. Kroz ostatak godine, cijene su naglo padale do ispod $40 u prosincu. Nakon još jednog reza OPEC-a za 4.2 milijuna barela dnevno u siječnju 2009. godine, cijene su počele polako rasti. U veljači 2011. godine cijene su porasle kao posljedica pada izvoza Libije za vrijeme građanskog rata.

Posljednji period od pet godina cijena nafte prikazan je na sljedećem linijskom grafikonu. Možemo primjetit nagli pad cijena, koje su se do lipnja 2014. godine, kada su dosegle razinu od preko $146, držalne na stabilnoj razini. Ovaj pad je dotakao dno u veljači 2016. godine ispod $34. Ovdje je riječ o NYMEX WTI futures ugovorima prikazanih na grafikonu u kontinuiranom obliku.



Slika 1: Futures WTI CL, 2012 – 2016. Interactive brokers, NYMEX podaci

Izvor: Prikaz autora

2.2.Vrste nafte

Proizvodima sirove nafte trguje se na različitim burzama. Trgovina može biti likvidna samo u standardiziranim proizvodima, zbog čega se standardizacija temelji na različitim karakteristikama među kojima ubrajamo: kemijske profile, lokacije dostave i financijske uvjete. Lokacije dostave su detaljnije opisane u sljedećem poglavlju.

Najvažniji kemijski profil je API gravitacija. API označava American Petroleum Institute, najveće udruženje trgovine naftom i zemnim plinom u SAD-u. API gravitacija je mjera težine ili lakoće nafte u odnosu na vodu, što predstavlja inverznu vrijednost u odnosu na gustoću naspram vode. U slučaju da je veća od 10, nafta je lakša od vode i pluta na njoj. Ispod 10 je teža i tone u vodi. Nafta se smatra teškom ako ima API od 30 ili niže, srednjom ako je API između 31 i 33, te lakom ako je API 34 ili viša. U pravilu se smatra da je nafta vrijednija ako ima visoku API gravitaciju. Druga najvažnija kemijska karakteristika je sadržaj sumpora. Nafta je slatka ako je sadržaj sumpora manji od 0.5% težine te kisela ako je sadržaj sumpora veči od 1% težine. WTI Crude je lagan (visok API) i sladak (mali udio sumpora), što ga čini idealnim za rafiniranje u goriva motornih vozila. Brent nije lagan kao WTI ali se i dalje smatra naftom visoke kvalitete. OPEC košarica je teža i kiselija u odnosu na Brent.

Financijski uvjeti se najčešće (vrijedi za WTI i Brent) odnose na cjenovne kvote koje se baziraju na FOB (Free On Board) principu koji specifizira kada se vlasništvo dobara transferira sa prodavača na kupca. Prema FOB standardu, prodavatelj snosi troškove transporta dobara do lokacije dostave, kao i troškove tovarenja. Kupac snosi troškove osiguranja, odtovarenja i transporta do krajnje destinacije. Najvažniji detalj standarda je prijenos rizika, koji se događa u trenutku kada su dobra utovarena na plovilo.

Najpoznatiji naftni proizvodi su WTI crude, *North Sea Brent* crude (skraćeno Brent) i UAE Dubai crude. Kako su u sljedećem poglavlju pobliže opisani WTI i Brent iz razloga što su najdominantniji u svijetu te su korišteni za analizu u ovom radu, ovdje ćemo površno opisati dva alternativna mjerila, ORB i Dubai crude.

OPEC *Reference Basket* (ORB) je vagani prosjek naftne sirovine proizvedene od strane članica OPEC kartela. Treba uzeti u obzir da, u usporedbi sa WTI i Brentom, ORB košarica je teža, zbog čega je povijesno cijena barela bila niža u odnosu na navedene. Također treba uzeti u obzir da je OPEC kartel kroz povijest aktivno regulirao cijenom upravo ovog mjerila nafte. ORB uključuje sljedeće naftne sirovine: Saharan Blend (Algeria), Girassol (Angola), Oriente (Ecuador), Rabi Light (Gabon), Minas (Indonesia), Iran Heavy (Iran), Basra Light (Iraq), Kuwait Export (Kuwait), Es Sider (Lybia), Bonny Light (Nigeria), Quatar Marine (Quatar), Arab Light (Saudi Arabia), Murban (UAE) i Merey (Venezuela).

Dubai crude je kiseo i srednje težine, ekstrahira se u Dubaiju. Iako je slabije kvalitete u odnosu na ostale, često se koristi kao globalno mjerilo iz razloga što je jedan od rijetkih naftnih sirovina koje se vade u Perzijskom zaljevu, te je eksport pretežito usmjeren prema Aziji.

2.3.WTI i Brent crude

Brent predstavlja globalno mjerilo za sirovu naftu općenito u Atlantskom oceanu. Također je konkretno mjerilo za naftu sa niskom sadržajem sumpora, popularno zvanom „slatkom“. U sektoru Ujedinjenog Kraljevstva Sjevernog mora, nalazi se Brent polje koje se dostavlja kroz cjevovod do terminala pod nazivom Sullom Voe u Ujedinjenom Kraljevstvu. Navedeno je prektički sinonim za futures i opcijske ugovore Brent nafte. Kao baza za cijene futures i opcijskih ugovora Brent-a koristi se definirana cijena na tržištu plovnih tereta nafte, gdje nafta dolazi u diskretnim količinama. Plovila dostavljaju teret nafte sa Sjevernog mora na različite lokacije. Nafta se vadi na više mjesta na Sjevernom moru, po ćemu je dobila kraticu koja označava pojedina mjesta: BFOE (*Brent-Forties-Oseberg-Ekofisk* obitelj sirove nafte Sjevernog mora). Ovo tržište, koje predstavlja bazu za definiranje cijene 65% svjetske trgovine u sirovoj nafti, nije samo po sebi spot tržište. To je zapravo kratkoročno forward tržište pod utjecajem CFD (*contract for difference*) ugovora, koji su derivirani iz forward krivulje (krivulja na temelju različitih isticanja ugovora) Brent futures ugovora i kratkoročnih opcija. Brent futures ugovorima se trguje na burzama ICE i NYMEX.

WTI je kontinentalni cjevovod gdje se sirova nafta kontinuirano prenosi u gotovo konstantnim količinama. Industrija sirove nafte u SAD-u započela je u zapadnoj Pensilvaniji i Istočnom Ohiu te se potom preselila otkrivanjem Meksičkog Zaljeva, Oklahome te Zapadnog i Istočnog Teksasa. U Kanadi je započela u južnom Ontariu te se potom preselila u Albertu. Oklahoma je ključna iz razloga što je bilo potrebno izgraditi dalekosežne cjevovode prema centrima za refiniranje u „Midwestu“. Tada nastaje terminal cjevovoda u Cushingu. Kako je Cushing postao sjecište preko kojeg se sirova nafta kreće iz svih navedenih lokacija u centre za refiniranje, Cushing je postao mjerilo za cjevovodnu sirovu naftu u SAD-u. Cijena na ovom tržištu definira se u vremenskom razdbolju između 25. dana u mjesecu i prvog dana u sljedećem mjesecu. Tada se po definiranoj fiksnoj cijeni, nafta srazmjerno prenosi cjevovodima u sljedećem mjesecu.

2.4. Futures ugovori nafte

Futures ugovori su proizvodi kojima se organizirano trguje na velikim svjetskim burzama. Futures ugovori nafte su najpopularniji među svim ugovorima sirovina. Postoje dvije velike burze gdje se gotovo sav volumen trgovanja odvija, *CME(Chicago Mercanite Exchange)* i *ICE(Intercontinental Exchange). Chicago Mercanite Exchange* upravlja prosječno sa preko tri milijarde ugovora godišnje. Sastoji se od više burzi: *CME, CBOT, NYMEX i COMEX*, gdje se trguje futuresima, opcijama i drugim proizvodima. *Intercontinental Exchange* pokrenuta je 2000. godine te uz globalne futures proizvode također pruža i OTC usluge, također je u brzom porastu volumena trgovanja.

Na *CME* burzi izlistano je preko stotinu različitih proizvoda koji spadaju u subkategoriju sirove nafte. Mnogi od tih proizvoda imaju gotovo nepostojeći volumen trgovanja. Prema volumenu možemo vidjeti da je *WTI* na prvom mjestu, nakon kojeg slijede *LSS, WTI Midland, Canadian Light Sweet Oil i Dubai crude*. U odnosu na ostale, *WTI* ima značajno veći interes te ćemo specifikacije ugovora pobliže opisati.

Jedan *WTI* crude oil futures ugovor na *CME* burzi se odnosi na 1000 barela. Iskazana cijena je u američkim dolarima te je minimalna cjenovna fluktuacija $0.01 po barelu, što iznosi $10 po ugovoru. Izlistani ugovori se protežu do devet godina u budućnost na način da je svaki mjesec izlistan za trenutnu godinu i sljedećih pet, dok su samo lipanj i prosinac izlistani za godine nakon šeste. Pri isticanju pojedinog ugovora, 1000 barela nafte se fizički dostavljaju imatelju ugovora. Iako se preko 99% ugovora zatvara prije isticanja, što rezultira samo novčanim profitkom ili gubitkom, mali postotak koji istekne ima točno određenu dostavu barela. Temeljena je na spomenutoj *FOB* metodi na bilo kojem cjevovodu ili skladištu u Crushingu, Oklahomi. Dodatni detalji su specifizirani na web stranicama CME burze koje nije potrebno daljnje analizirati u ovom radu.

ICE burza također ima izlistan veliki broj energetskih i drugih proizvoda od kojih Brent crude ima najveći volumen trgovanja. Jedan ugovor dostavlja 1000 barela. Cijena je također izlistana u američkim dolarima, $0.01 po barelu. Glavna razlika u odnosu *WTI* naspram *CME*-u je što se po isteku ugovora ne dostavlja fizički proizvod, već protuvrijednost ugovora u američkim dolarima, odnosno imatelju ugovora na datum isteka se računa profit ili gubitak, pripisuju na stanje računa i zatvara pozicija. Nikakav prijenos barela se ne događa, što mnogo olakšava proces.

Jedna od ključnih karakteristika zbog kojih je trgovina ovim ugovorima popularna je visoka financijska poluga. Ako primjerice puna vrijednost jednog ugovora pri cijeni od $50 po barelu iznosi $50 000, imatelj ugovora ne treba platiti puni iznos, već staviti marginalni iznos definiran od strane burze i/ili brokera. Ovaj iznos se najčešće broji u nekoliko tisuća američkih dolara, a i manje ako se trguje u vremenskim intervalima manjima od jednog dana.

2.5. Odabir podataka

U izboru između *WTI* i *Brent* futures ugovora odabrali smo *Brent crude*. Obje vremenske serije bi u pravilu trebale biti pod istim utjecajem eksternih varijabli te imati gotovo jednake karakteristike vremenskih serija kao što su stacionarnost, sezonalnost i trend komponente. Kako se dvije trećine svjetske trgovine bazira na cijeni *Brent crude* globalnog mjerila, odabrali smo ovaj proizvod. Dnevne i tjedne cijene i volumeni trgovanja povučeni su sa *Thomson Reuters* baze podataka.

3. MODELIRANJE VREMENSKIH SERIJA

3.1.Testiranje stacionarnosti i odabir razdoblja

Odabrana su dva razdoblja: petogodišnje u rasponu od 31.5.2006. do 25.5.2010. godine te desetogodišnje u rasponu od 2.6.2005. do 22.5.2015. godine. Prvi set podataka odnosi se na dnevne cijene zatvaranja (*closing price*), dok se drugi odnosi na tjedne cijene zatvaranja. Prvi set sadrži 1288 opservacija, dok drugi sadrži 469. Različiti uzorci uzeti su u svrhu usporedbe i testiranja konzistentnosti statističkih metoda kroz različite setove podataka. Isto tako, moguće su bolje performanse pojedinih modela za različito vremensko razdoblje. Koristit će se statistički alat Rstudio (radna okolina programskog jezika R).

Da bi se bolje upoznali sa našim vremenskim nizovima navedena je osnovna deskriptivna statistika na sljedećem prikazu.

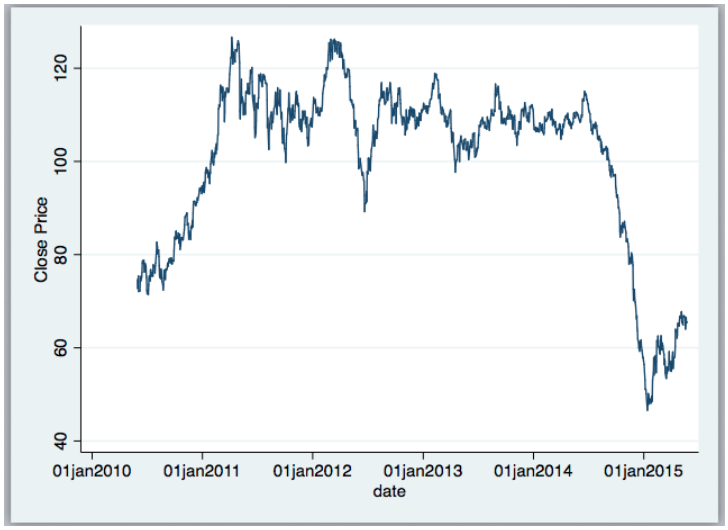
```
> summary(timeseries5y)
     Index              timeseries5y
 Min.   :2010-06-01   Min.   : 46.59
 1st Qu.:2011-04-18   1st Qu.: 93.62
 Median :2012-03-05   Median :107.78
 Mean   :2012-03-05   Mean   :100.62
 3rd Qu.:2013-01-21   3rd Qu.:111.57
 Max.   :2013-12-09   Max.   :126.65
> summary(timeseries10y)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  38.37   70.62   92.46   89.91  109.60  144.50
> sd(timeseries5y)
[1] 17.90388
> sd(timeseries10y)
[1] 23.09216
```

Slika 2: Deskriptivna statistika za oba vremenska niza

Izvor: Prikaz autora

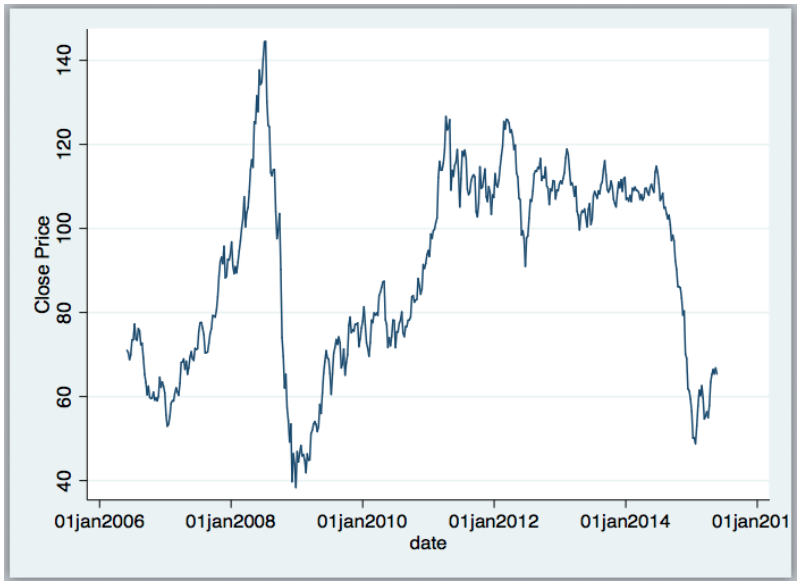
Osim standardnih vrijednosti minimuma, maximuma, mediana, moda i kvartila, izračunata je i standardna devijacija niza. Prvim pogledom može se primjetiti da vremenski niz dužeg vremenskog razdoblja ima i veću standardnu devijaciju. Intuitivno objašnjenje je uključenost financijske krize 2008. godine u ovo razdoblje.

Spomenuti geopolitički događaji bili su manje značajni u posljednjih 5 godina (iako i dalje prisutni) što predstavlja dodatnu pogodnost za modeliranje. Osim događaja u Libiji i nagli pad cijena 2014. godine uvjetovan fundamentalnim razlozima, među kojima su: nakupljanje zaliha, akcije OPEC-a i funkcija potražnje. Drugi vremenski set podataka ima veći broj geopolitičkih utjecaja. Od značajnog utjecaja je kreditna financijska kriza 2008. godine.



Slika 3: Linijski graf - dnevne opservacije cijena zatvaranja Brent futures ugovora, 31.5.2006. - 25.5.2010. godine

Izvor: Prikaz autora



Slika 4: Linijski graf - tjedneopservacije cijena zatvaranja Brent futures ugovora, 2.6.2005. - 22.5.2015

Izvor: Prikaz autora

Prvim korakom provjeravamo potrebu za logaritmiranjem serija. Često se vremenske serije cijena modeliraju korištenjem logaritmiranih vrijednosti povrata, odnosno razlike između trenutnih i prethodnih t-1 log cijena. Lutkepohl i Xu (2009) pokazali su da logaritmiranjem vremenskih nizova u slučaju svjetskih dioničkih indeksa rezultira AR(1) modelom sa manjim *MSE* (*Mean Squared Error*) vrijednostima u većini vremenskih perioda. Također su pronašli da korištenjem logartmiranih vrijednosti u slučajima kada se bezuspješno stabilizira varijanca može biti i pogubno za prognostičke modele. Treba uzeti u obzir da se vremenski niz potrošnih dobara, kao što je vremenski niz sirove nafte, znatno razlikuje od dioničkih indeksa, koji najčešće imaju dugoročnu tendenciju blagog rasta. Cilj logaritmiranja je stabilizacija varijance te približavanje *Gaussovom* procesu. Za provjeru potrebe logaritmiranja koristit će se *Box-Cox* transformacija.

Box-Cox transformacija, transformira svaku varijablu y u y^{λ} gdje vrijedi:

$$\frac{y^{\lambda}-1}{\lambda}, ako\ \lambda\ \neq 0 \qquad (1)$$

$$\log(y), ako\ \lambda = 0 \qquad (2)$$

Guerrerovom (1993) metodom, gdje lambda minimizira koeficijent varijacije serije, dobiva se lambda vrijednost blizu 1 za obje serije što upućuje da transformacija nije potrebna. Dodatno, usporedbom *Augmented Dickey Fuller* testova za logaritmirane i nelogaritmirane vrijednosti obje serije, p-vrijednost logaritmiranog niza sa dnevnim opservacijama (0.6749) je gotovo jednaka u odnosu na vrijednost polaznog niza sa dnevnim opservacijama (0.6873), a manja (0.2114 za logaritmirani niz u odnosu na 0.3588 za originalni), iako je statistički značajna u seriji sa dnevnim opservacijama.

Nadalje, za provjeru stacionarnosti vremenskih nizova te s tim i potrebu za diferenciranjem, koristimo se spomenutim *Augmented Dickey Fuller* testom jediničnog korjena. Nestacionarnost je jasno vidljiva i sa linijskih grafičkih prikaza.

Na rezultatima je vidljiva nestacionarnost oba vremenska niza. p-vrijednost iznosi 0.6873 u dnevnom vremenskom nizu, dok u tjednom iznosi 0.3588, što je više u odnosu na razine signifikatnosti od 1%, 5% i 10%, što znači da ne možemo odbaciti nultu hipotezu da vremenski nizovi prate proces jediničnog korjena. Stoga zaključujemo da vremenske serije treba diferencirati.

```
> adf.test(timeseries5y)

        Augmented Dickey-Fuller Test

data:  timeseries5y
Dickey-Fuller = -1.7425, Lag order = 10, p-value = 0.6873
alternative hypothesis: stationary

> adf.test(timeseries10y)

        Augmented Dickey-Fuller Test

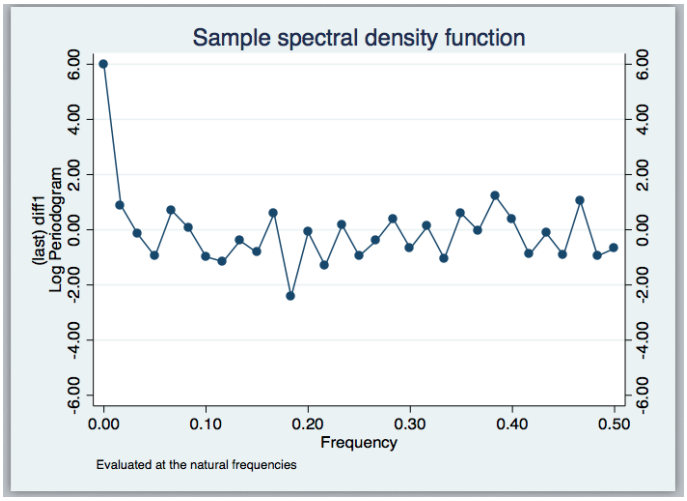
data:  timeseries10y
Dickey-Fuller = -2.5181, Lag order = 7, p-value = 0.3588
alternative hypothesis: stationary
```

Slika 5: Output Augmented Dickey-Fuller testa u statistčkom alatu R

Izvor: Prikaz autora

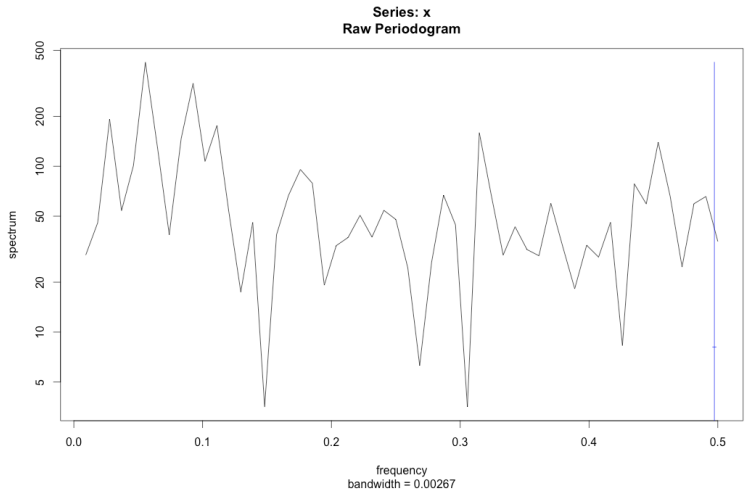
3.2. Sezonalnost

Transformiranjem vremenskih nizova u mjesečne opservacije otvaramo put za jednostavniju interpretaciju periodograma. Zadnja cijena zatvaranja svih opservacija u mjesecu uzeta je kako bi predstavljala mjesečne opservacije umjesto prosječne. To je standardna notacija izvještavanja u financijskoj industriji. Periodogram se koristi za identificiranje dominantnih perioda frekvencija vremenskih serija. Svaki vremenski niz može se prikazati kao kombinacija valova kosinusa (i sinusa) sa periodima koji se međusobno razlikuju. Promatramo vrijeme koje je potrebno za završiti cijeli ciklus i amplitudu, odnosno maksimalne i minimalne vrijednosti unutar pojedinog ciklusa. Nakon uzimanja prve diferencije, periodogram ne prikazuje značajnu sezonalnosti u pojedinom ciklusu. Ovaj zaključak vrijedi za oba vremenska niza.



Slika 6: Periodogram – differencirani niz mjesečnih opservacija; 31.5.2006. do 25.5.2010.

Izvor: Prikaz autora



Slika 7: Periodgram – diferencirani niz mjesečnih opservacija; 2.6.2005. do 22.5.2015.

Izvor: Prikaz autora

3.3. Trend komponente

Na temelju zaključaka koje je donio Perron (1988) u radu "*The Great Crash, the Oil Price Shock and the Unit Root Hypothesis*" zaključujemo da strukturni lomovi u vremenskim serijama, kao primjerice geoplitički utjecaji u svijetu naftne industrije, imaju tendenciju utjecati na rezultate testova stacionarnosti. Uvedemo li strukturne promjene na najznačajnije događaje, koji imaju ekstreman utjecaj na razinu cijena, možemo provjeriti ima li značajnih razlika u rezultatima.

Zivot i Andrews (1992) razvili su test koji koristi *dummy* varijable za različite strukturne lomove pomoću kojeg možemo donijeti zaključke o stacionarnosti niza. Test dopušta po jedan lom u *interceptu* i trendu vremenske serije. Predlažu intervencijski model za razliku od aditivnog modela kojeg je koristio Perron. Regresijska formula glasi:

$$\hat{y}_t^B = \hat{a}^B \hat{y}_{t-1}^B + \sum_{i=1}^{k} \hat{c}_i^B \, \Delta\hat{y}_{t-1} \hat{e}_t \qquad (3)$$

Za koju vrijedi sljedeće: $\hat{y}_t^B$ su reziduali regresije sa y_t te varijable koje objašnjavaju reziduale sadrže konstantan vremenski trend. Strukturni lom se tretira kao endogeni događaj.

Nulta hipoteza se prezentira u obliku:

$$y_t = a_0 + y_{t-1} + e_t \quad (4)$$

Zivot Andrews testom nad petogodišnjem vremenskom nizu možemo vidjeti da uz endogeno odvojeni strukturni lom iz vremenske serije u opservaciji broj 421, t-vrijednost od -3.6037 viša je u odnosu na ponuđene razine signifikatnosti, što znači da ne možemo odbaciti nultu hipotezu da vremenski niz prati proces jediničnog korjena. Uz endogeni strukturni lom, vremenska serija i dalje nije stacionarna što potvrđuje potrebu za diferenciranjem niza. Jednaki rezultati javljaju se u slučaju desetogodišnjeg vremenskog niza. Uz predloženi strukturni lom u opservaciji broj 1103, t-vrijednost je viša u odnosu na kritične vrijednosti od 1%, 5% i 10%. Ne možemo odbaciti nultu hipotezu, stoga modeliranje strukturnog loma ne dodaje vrijednost.

```
Title:
 Zivot & Andrews Unit Root Test

Test Results:

  Residuals:
      Min       1Q   Median       3Q      Max
-16.2104  -1.5683   0.2164   2.1169  11.4893

  Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.758729   0.701089   2.509  0.01247 *
y.l1         0.974746   0.009078 107.373  < 2e-16 ***
trend        0.003062   0.001795   1.706  0.08869 .
y.dl1        0.018468   0.046225   0.400  0.68969
y.dl2        0.085000   0.046243   1.838  0.06669 .
du          -2.087546   0.739195  -2.824  0.00495 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

  Residual standard error: 3.582 on 460 degrees of freedom
    (3 observations deleted due to missingness)
  Multiple R-squared:  0.9762,  Adjusted R-squared:  0.976
  F-statistic:  3779 on 5 and 460 DF,  p-value: < 2.2e-16


  Teststatistic: -2.7819
  Critical values: 0.01= -5.34 0.05= -4.8 0.1= -4.58

  Potential break point at position: 421
```

```
Title:
 Zivot & Andrews Unit Root Test

Test Results:

  Residuals:
      Min       1Q   Median       3Q      Max
-10.3708  -0.7793  -0.0176   0.8649   6.3567

  Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.3971034  0.3656623   3.821 0.000139 ***
y.l1         0.9869464  0.0036223 272.466  < 2e-16 ***
trend        0.0000275  0.0001515   0.182 0.855976
y.dl1       -0.0570991  0.0278139  -2.053 0.040286 *
y.dl2       -0.0325884  0.0277930  -1.173 0.241198
du          -0.7460232  0.2246979  -3.320 0.000925 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

  Residual standard error: 1.495 on 1279 degrees of freedom
    (3 observations deleted due to missingness)
  Multiple R-squared:  0.993,   Adjusted R-squared:  0.993
  F-statistic: 3.648e+04 on 5 and 1279 DF,  p-value: < 2.2e-16


  Teststatistic: -3.6037
  Critical values: 0.01= -5.34 0.05= -4.8 0.1= -4.58

  Potential break point at position: 1103
```

Slika 8: Zivot-Andrews test jediničnog korjena za 10 godišnji vremenski niz (lijevo) i 5 godišnji vremenski niz (desno)

Izvor: Prikaz autora

3.4. AR model

AR (*Autoregressive*) model definira da output varijabla ovisi o linearnosti svojih prethodnih (t-1) vrijednosti te o stohastičkoj vrijednosti. AR je najjednostavniji model koji će se razraditi te predstavlja komponentu ARIMA modela koji se definira u nastavku. Formula koja opisuje AR(p) model je sljedeća:

$$Y_t = \alpha + \beta Y_{t-1} + \varepsilon_t \quad (5)$$

AR(1) model primjenjujemo na diferencirani vremenski niz. Rezultati modeliranja diferenciranih nizova AR(1) vidljivi su na sljedećem prikazu.

```
> summary(arimasimple5ydiff)

Call:
arima(x = diff(timeseries5y), order = c(1, 0, 0))

Coefficients:
          ar1  intercept
      -0.0552    -0.0071
s.e.   0.0278     0.0397

sigma^2 estimated as 2.258:  log likelihood = -2350.42,  aic = 4706.84

Training set error measures:
                      ME     RMSE      MAE MPE MAPE      MASE         ACF1
Training set -1.132898e-05 1.502814 1.103259 NaN  Inf 0.6701169 -0.001671267
```

```
> summary(arimasimple10ydiff)

Call:
arima(x = diff(timeseries10y), order = c(1, 0, 0))

Coefficients:
         ar1  intercept
      0.0180    -0.0116
s.e.  0.0462     0.1698

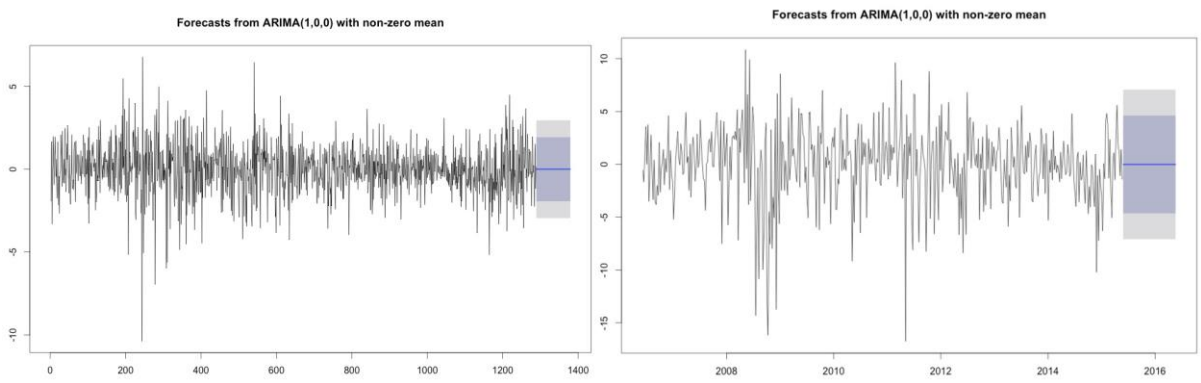
sigma^2 estimated as 13.02:  log likelihood = -1264.63,  aic = 2535.25

Training set error measures:
                       ME     RMSE      MAE  MPE MAPE      MASE         ACF1
Training set -0.0005105974 3.608363 2.632158 -Inf  Inf 0.6912513 -0.001418766
```

Slika 9: AR(1) diferenciranih vremenskih nizova: niz 5 godina (lijevo), niz 10 godina (desno)

Izvor: Prikaz autora

Na sljedećim grafičkim prikazima možemo vidjeti predviđene vrijednosti AR(1) modela. Predviđena je 91 opservacija (jedan kvartal) za vremenski niz od 5 godina te 52 opservacije (1 godina) za vremenski niz od 10 godina. Treba obratiti pažnju na obojane intervale signifikatnosti od 80% i 95%. Kako se radi o diferenciranim vrijednostima, grafički prikazi odnose se na apsoultne dnevne (u prvom slučaju) i tjedne (u drugom)povrate.



Slika 10: Grafički prikaz AR(1) prognoze vremenskih nizova: 91 za 5 godišnji niz (lijevo), 52 za 10 godišnji (desno)

Izvor: Prikaz autora

Ljung-Box testom provjeravamo autokorelacije reziduala našeg jednostavnog ARIMA(1,0,0) modela. Formalna funkcija statističke vrijednosti testa glasi:

$$Q = n(n+2)\sum_{k=1}^{h}\frac{\hat{p}_k^2}{n-k} \quad (6)$$

Objašnjenje oznaka:

n - veličina uzorka

$\hat{p}_k^2$ - autokorelacija uzorka na lagu k

h- broj lagova koji se testira

Visoke p-vrijednosi testa u obje serije pokazuju da residuali prate *white noise* proces.

```
> Box.test(residuals(arimasimple10ydiff, type="Ljung", fitdf=1), lag=24)

        Box-Pierce test

data:  residuals(arimasimple10ydiff, type = "Ljung", fitdf = 1)
X-squared = 31.535, df = 24, p-value = 0.1389

> Box.test(residuals(arimasimple5ydiff, type="Ljung", fitdf=1), lag=24)

        Box-Pierce test

data:  residuals(arimasimple5ydiff, type = "Ljung", fitdf = 1)
X-squared = 14.531, df = 24, p-value = 0.9337
```

Slika 11: Ljung-Box test autokorelacije reziduala AR(1) modela

Izvor: Prikaz autora

3.5. ARIMA model

ARMA (Autoregressive Moving Average) (p,q) model je vremenskog niza Y_t koji ima sljedeći oblik:

$$Y_t = \phi_0 + \sum_{i=1}^{p} \phi_1 Y_{t-1} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \epsilon_t \quad (7)$$

Objašnjenje oznaka:

p - broj izraza autoregresije

q - broj lagiranih reziduala

ϕ - koeficijenti izraza autoregresije i konstanta

θ - koeficijenti izraza pomičnog prosjeka

Vremenski niz je u ARIMA (*Autoregressive Integrated Moving Average*) formatu ako $\nabla^d Y_t$ je stacionarni ARMA proces. Oznaka d odnosi se na broj difernecije koji se uzima originalnom vremenskom nizu.

ARIMA modeliranje je provedeno u statističkom programskom jeziku R. Korištenjem funkcije *auto.arima()* definiran je najbolji ARIMA model. Funkcija koristi varijaciju *Hyndman i Khandakar* algoritma koji kombinira test jediničnog korjena, minimiziranje AICc i MLE za dobivanje optimalnog ARIMA modela. Broj potrebnih diferencija d dobiva se automatiziranim ponavljanjem KPSS testova. Kako smo već definirali potrebu za

diferenciranjem, rezultati nam mogu poslužiti kao potvrda prethodnih testova. Vrijednosti *p* i *q* biraju se minimiziranjem AICc-a nakon diferenciranja serije *d* broj puta. BIC (*Bayesian Information Criterion*) je također korišten.

AIC ćemo pobliže proučiti, jer je temeljni kreterij odabira modela algoritma. *Akaikeov informacijski kriterij* je mjera relativno kvalitete i konzistentnosti statističkih modela za dani set podataka. AIC vrijednost testiranog modela se dobiva na sljedeći način:

$$AIC = 2k - 2\ln(L) \quad (8)$$

Objašnjenje oznaka:

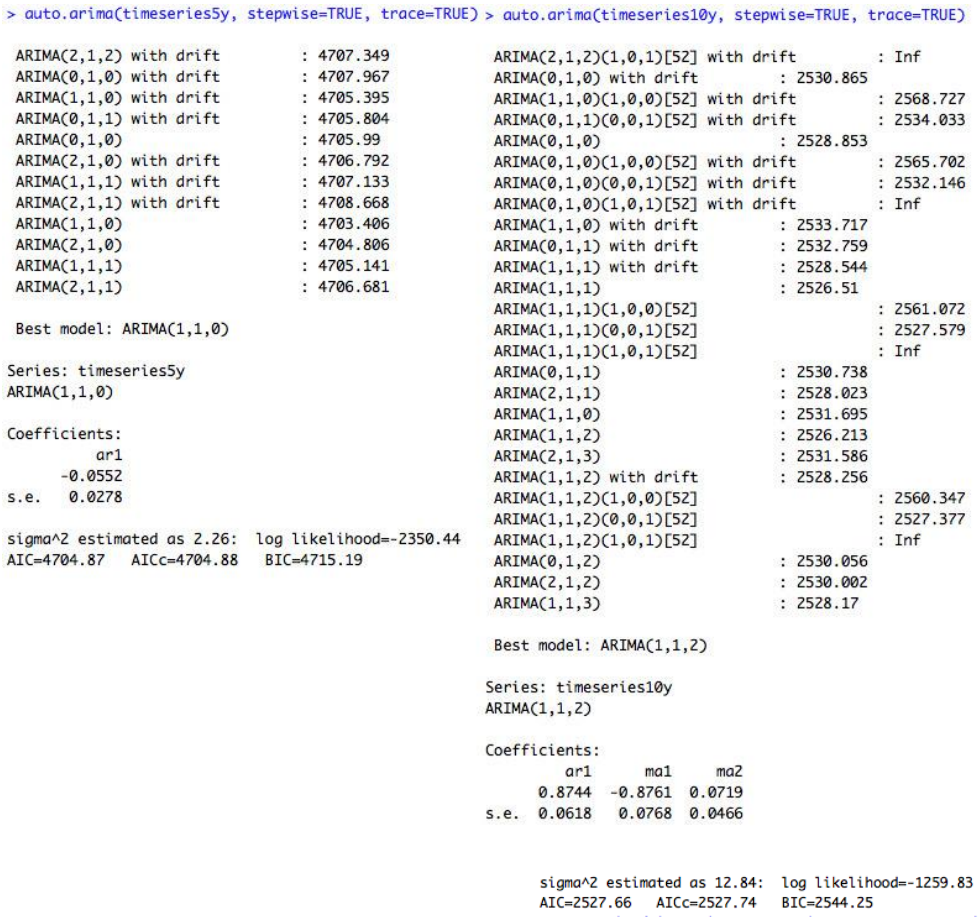
k - procjenjeni broj parametara u modelu

L - maximalna vrijednost funkcije vjerojatnosti modela

AICc dodaje ispravak za konačne uzorke. Uz pretpostavku da je model linearan te da ima normalno distribuirane reziduale, formula za AICc je sljedeća:

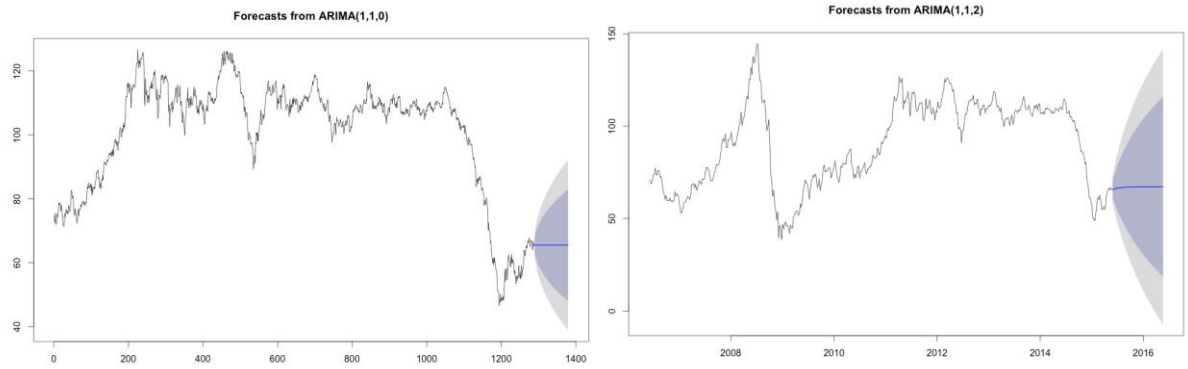
$$AICc = AIC + \frac{2k(k+1)}{n-k-1} \quad (9)$$

Vrijednosti za testirane modele i odabir modela algoritmom prikazni su na sljedećoj slici. Potom možemo vidjeti grafički prikaz prognoziranih vrijednosti.



Slika 12: Rezultati filtriranja ARIMA modela na temelju informacijskih kriterija: vremenski niz sa dnevnim opservacijama (lijevo), vremenski niz sa tjednim opservacija (desno)

Izvor: Prikaz autora

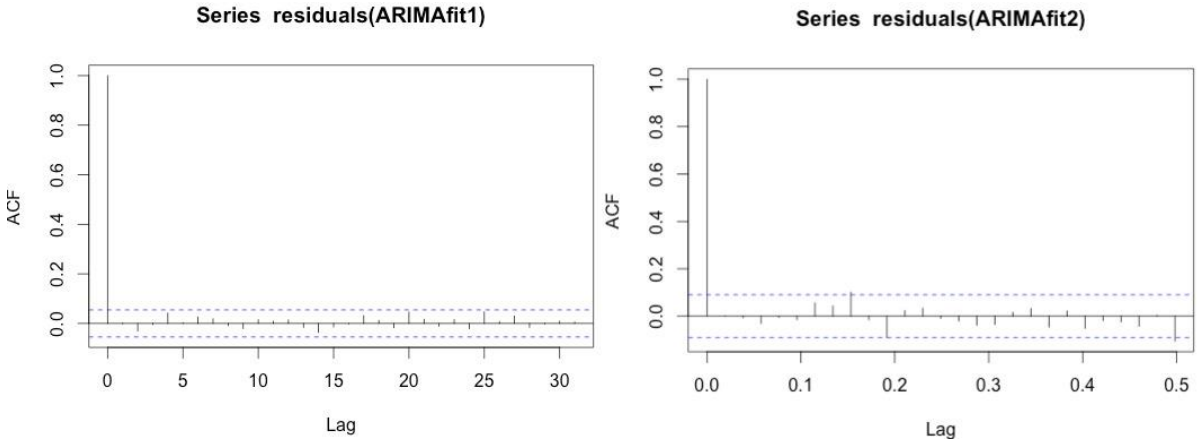


Slika 13: Prognozirane vrijednosti vremenskih nizova: 1 kvartal (91 opservacija) za vremenski niz s dnevnim opservacijama (lijevo), 1 godina (52 opservacije) za vremenski niz sa tjednim opservacija

Izvor: Prikaz autora

Vidimo da *auto.arima()* funkcija bira ARIMA(1,1,0) za vremenski niz dnevnih opservacija te ARIMA(1,1,2) za vremenski niz tjednih opservacija. Razine pouzdanosti od 80% i 95% su istaknute u grafičkim prikazima. Odmah je vidljivo da same predviđene opservacije nemaju visoku prognostičku vrijednost, ali sami intervali pouzdanosti mogu dati potencijalno korisnu informaciju opsega kretanja vremenskog niza u definiranom vremenu.

Visoka p-vrijednost *Ljung-Box* testova i ACF grafovi reziduala ARIMA modela potvrđuju da reziduali prate *white noise* proces.



Slika 14: Autokorelacijske funkcije reziduala ARIMA (1,1,0) modela (lijevo) te ARIMA (1,1,2) modela (desno)

Izvor: Prikaz autora

```
> Box.test(residuals(ARIMAfit1), type="Ljung", fitdf=1, lag=24)

        Box-Ljung test

data:  residuals(ARIMAfit1)
X-squared = 14.694, df = 23, p-value = 0.9053

> Box.test(residuals(ARIMAfit2), type="Ljung", fitdf=3, lag=24)

        Box-Ljung test

data:  residuals(ARIMAfit2)
X-squared = 18.777, df = 21, p-value = 0.5994
```

Slika 15: Ljung-Box test autokorelacije reziduala ARIMA(1,1,0) modela (gore) te ARIMA(1,1,2) modela (dolje)

Izvor: Prikaz autora

3.6. GARCH model

Veliki broj radova bavi se problematikom modeliranja i predviđanja volatilnosti nafte, naročito modeli vrste GARCH. Primjerice, Kang et al. (2009) daju prednost frakcijski integriranom GARCH modelu (FIGARCH). Cheong (2009) zaključuje da bi standardni GARCH i FIAPARCH pokazali bolje rezultate. Mohammadi i Su (2010) daju prednost APARCH modelima. Wei et al. (2010) zaključuju da se od devet različitih GARCH modela niti jedan nije istaknuo u prognoziranju vremenskih nizova. U većini radova GARCH modeli se grade na temelju AR(1) procesa. U ovom radu će se također graditi na prethodno definiranim optimalnim ARIMA modelima: ARIMA(1,1,0) i ARIMA(1,1,2).

Najpopularniji model u literaturi je linearni GARCH (*General Autoregressive Conditional Heteroskedasticity*). GARCH ima sposobnost modeliranja grupiranja volatilnosti, jednu od karakteristika cijene sirove nafte. Uveo ga je Bollersev (1986). Model je temeljen na ARCH modelu, kojeg je prvi put predložio Engle (1982). ARCH model polazi od postavke da se varijanca sadašnje inovacije temelji na inovacijama iz prethodnih vremenskih perioda. ARCH(q) model foruma je sljedeća:

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-1}^2 , \omega > 0, \alpha_i \geq 0, i > 0 \quad (10)$$

GARCH(p,q) model prikazuje se sljedećim izrazom:

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-1}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-1}^2, \omega > 0, \alpha_i \geq 0, \beta_i > 0, i > 0 \quad (11)$$

Uz uvjet da

$$\sum_{i=1}^{q} \alpha_i + \sum_{i=1}^{p} \beta_i < 0$$

Ako $\alpha + \beta = 1$, GARCH(1,1) reducira se u IGARCH(1,1) kojeg su predložili Engle i Bollerslev (1986). Ako je $p = 0$, GARCH se reducira u ARCH.

GARCH(1,1) koristi se za modeliranje AR(1) procesa u slučaju oba vremenska niza, ARIMA(1,1,0) u slučaju vremenskog niza sa dnevnim opservacija te ARIMA(1,1,2) u slučaju vremenskog niza sa tjednim opservacijama.

Od četiri dobivena modela, proučavaju se rezultati IGARCH(1,1) primjenjen na ARIMA(1,1,0) modelu, odnosno GARCH model na temelju optimalnog ARIMA modela dobivenog *Hyndman i Khandakar* algoritmom vremenske serije dnevnih opservacija za petogodišnje razdoblje. Usporeba svih modela nalazi se u zasebnom poglavlju.

```
*----------------------------------*
*          GARCH Model Fit         *
*----------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model      : sGARCH(1,1)
Mean Model       : ARFIMA(1,0,0)
Distribution     : norm

Optimal Parameters
------------------------------------
        Estimate  Std. Error   t value Pr(>|t|)
mu     -0.009491    0.035438  -0.26783 0.788829
ar1    -0.063479    0.028752  -2.20781 0.027258
omega   0.014436    0.006087   2.37172 0.017706
alpha1  0.033815    0.004966   6.80932 0.000000
beta1   0.960125    0.004960 193.58763 0.000000

Robust Standard Errors:
        Estimate  Std. Error   t value Pr(>|t|)
mu     -0.009491    0.038286  -0.24791 0.804206
ar1    -0.063479    0.029456  -2.15504 0.031159
omega   0.014436    0.007526   1.91811 0.055097
alpha1  0.033815    0.005148   6.56801 0.000000
beta1   0.960125    0.003111 308.66569 0.000000

LogLikelihood : -2283.455
```

```
Information Criteria
------------------------------------

Akaike       3.5563
Bayes        3.5763
Shibata      3.5562
Hannan-Quinn 3.5638

Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.2794  0.5971
Lag[2*(p+q)+(p+q)-1][2]    0.2983  0.9944
Lag[4*(p+q)+(p+q)-1][5]    0.8251  0.9695
d.o.f=1
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic   p-value
Lag[1]                     0.6753 0.4112087
Lag[2*(p+q)+(p+q)-1][5]   15.0563 0.0004292
Lag[4*(p+q)+(p+q)-1][9]   18.8812 0.0003881
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
            Statistic Shape Scale P-Value
ARCH Lag[3]     2.366 0.500 2.000  0.1240
ARCH Lag[5]     4.023 1.440 1.667  0.1718
ARCH Lag[7]     5.538 2.315 1.543  0.1755
```

Slika 16: Statistički pokazatelji za sGARCH(1,1) model primjenjen na ARIMA(1,1,0) za petogodišnju vremensku seriju dnevnih opservacija, prvi dio.

Izvor: Prikaz autora

```
Nyblom stability test
------------------------------------
Joint Statistic:  0.9234
Individual Statistics:
mu     0.52868
ar1    0.05754
omega  0.13826
alpha1 0.17160
beta1  0.16188

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:         1.28 1.47 1.88
Individual Statistic:    0.35 0.47 0.75

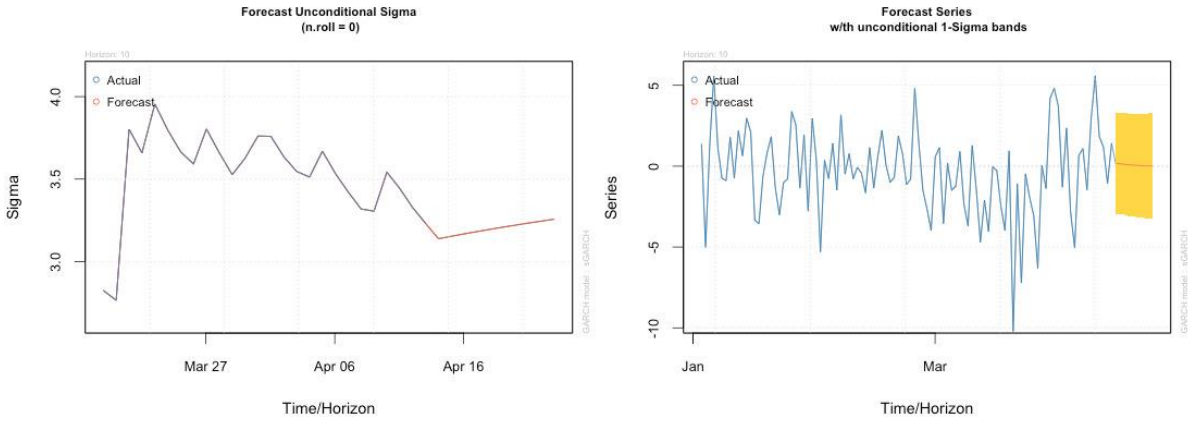
Sign Bias Test
------------------------------------
                   t-value    prob sig
Sign Bias           1.5566 0.11980
Negative Sign Bias  0.8352 0.40374
Positive Sign Bias  0.6581 0.51058
Joint Effect        6.7145 0.08158   *


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     43.96    0.0009578
2    30     50.09    0.0088403
3    40     62.65    0.0095186
4    50     73.33    0.0137486
```

Slika 17: Statistički pokazatelji za sGARCH(1,1) model primjenjen na ARIMA(1,1,0) za petogodišnju vremensku seriju dnevnih opservacija, drugi dio.

Izvor: Prikaz autora

Weighted Ljung-Box test pokazuje statistički značajnu p-vrijednost, što upućuje na nekoreliranost reziduala. p-vrijednost *Ljung-Box* testa za kvadrirane reziduale pri prvom *lagu* također upućuje na nekoreliranost. Možemo zaključiti da ARIMA proces odgovara vremenskom nizu. ARCH-LM test je oblik težinskog *portmanteau* testa za provjeru odgovaranja odabranog ARCH procesa. p-vrijednosti pri odabranim *lagovima* su visoke, što dovodi do zaključka da ARCH proces odgovara vremenskom nizu. *Nyblomov* test stabilnosti potvrđuje relativnu konzistentnost parametara kroz vrijeme.



Slika 18: Projecirane vrijednosti vremenske serije i standardne devijacije ARIMA(1,1,0)-GARCH(1,1) modela

Izvor: Prikaz autora

3.7. Varijacije GARCH modela

Razmatat ćemo tri alternativna GARCH modela.

APARCH model uključuje dodatak koji uračunava asimetrične utjecaje volatilnosti. Uključuje dva efekta: efekt poluge i *Taylor* efekt, koji pokazuje da je autokorelacija apsolutnih povrata uzorka najčešće veća u odnosu na kvadrirane povrate. Kako se u prethodnim radovima APARCH pokazao efikasnim, također ćemo ga razmatrati i ovdje. Definira se na sljedeći način:

$$\sigma_t^{\delta} = \omega + \sum_{i=1}^{q} \alpha_i \left(\epsilon_{t-1} + \gamma_i \epsilon_{t-1}\right)^{\delta} + \sum_{i=1}^{p} \beta_i \, \sigma_{t-1}^2, \omega > 0, \alpha_i \geq 0, \beta_i > 0, i > 0 \quad (12)$$

GJR-GARCH, model je razvijen od strane Glosten et al. (1993). Model nastoji uključiti potencijalno snažniji utjecaj negativnih šokova na volatilnost povrata. Varijanca GJR-GARCH(1,1) dobiva se sljedećom formulom:

$$\sigma_t^2 = \omega + [\alpha + \gamma D(x_{t-1} < 0)] x_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (13)$$

Gdje je D(.) indikatorska funkcija koja uzima vrijednost 1 ako je $x_{t-1} < 0$ što predstavlja scenarij loše vijesti te 0 za dobru vijest. Parametar γ određuje magnitudu asimetričnog utjecaja poluge.

Konačno, EGARCH je model predložen od strane Nelsona(1991) sa ciljem uključivanja asimetričnog odnosa povrata i volatilnosti. Izraz modela EGARCH(1,1) je:

$$\ln(\sigma_t^2) = \omega + \alpha e_{t-1} + \gamma(e_{t-1} - \mathrm{E}[e_{t-1}]) + \beta \ln(\sigma_{t-1}^2) \quad (14)$$

Gdje γ predstavlja parametar asimetrične poluge koji kvantificira obujam efekta volatilnosti poluge u iznosu magnitude α.

4. ANALIZA FUNDAMENTALNIH FAKTORA

4.1.Odabir fundalemntalnih faktora i njihovih vremenskih nizova

Odabrano je sedam nezavisnih varijabli za regresijski model. Pretpostavke utjecaja varijabli temeljene su na direktnoj povezanosti u slučaju sljedećih varijabli: prosječna dnevna potrošnja barela, prosječna dnevna proizvodnja barela, konačne zalihe u mjesecu,neto uvozi barela SAD-a te iskorištenost rafinerija (u postocima). Sljedeće dvije varijable odabrane su na temelju utjecaja globalizacije najznačajnijih svjetskih financijskih proizvoda: S&P index i Fed Funds kamatna stopa. Zbog ograničenosti dostupnosti dnevnih i tjednih vremenskih nizova mnogih od navedenih varijabli, regresija se temelji na mjesečnim podacima i to u razdoblju od lipnja 2006. do svibnja 2015. Pobliže ćemo opisati svaku od nezavisnih varijabli:

- **Prosječna dnevna proizvodnja barela:** kako znamo da dvije temeljne sile koja pokreću cijenu proizvoda su ponuda i potražnja, proizvodnja barela nafte bio je prvi logični odabir. U svom radu, Deng et al. (2012) proučili su utjecaje proizvodnje i drugih fundamentalnih faktora na cijenu sirove nafte. Treba uzeti u obzir da se ovdje odnosi na proizvodnju samo u SAD-u. Svjetsku proizvodnju mnogo je teže procjeniti, a proizvodnja SAD-a je vrlo precizno dokumentirana. Podaci su uzeti sa *U.S. Energy Information Administration*, državna administracija koja pobliže proučava i obavještava za sve energetske proizvode.

- **Prosječna dnevna potrošnja barela:** jednako kao i ostale varijable, podaci se odnose na SAD. Teoretsku pozadinu utjecaja možemo pronaći u radu Piroga (2005). Podaci su također uzeti sa službene stranice *EIA.*

- **Neto uvoz barela:** Naznačajniji uvoznih nafte u svijetu je SAD. Iako se ova varijabla može promatrati kao potražnja, neto uvoz definira potrebe za dodatnom naftom, iznad interne ponude što može doprinijeti objašnjavanju nezavisne varijable. Izvora podataka je službena stranica *EIA.*

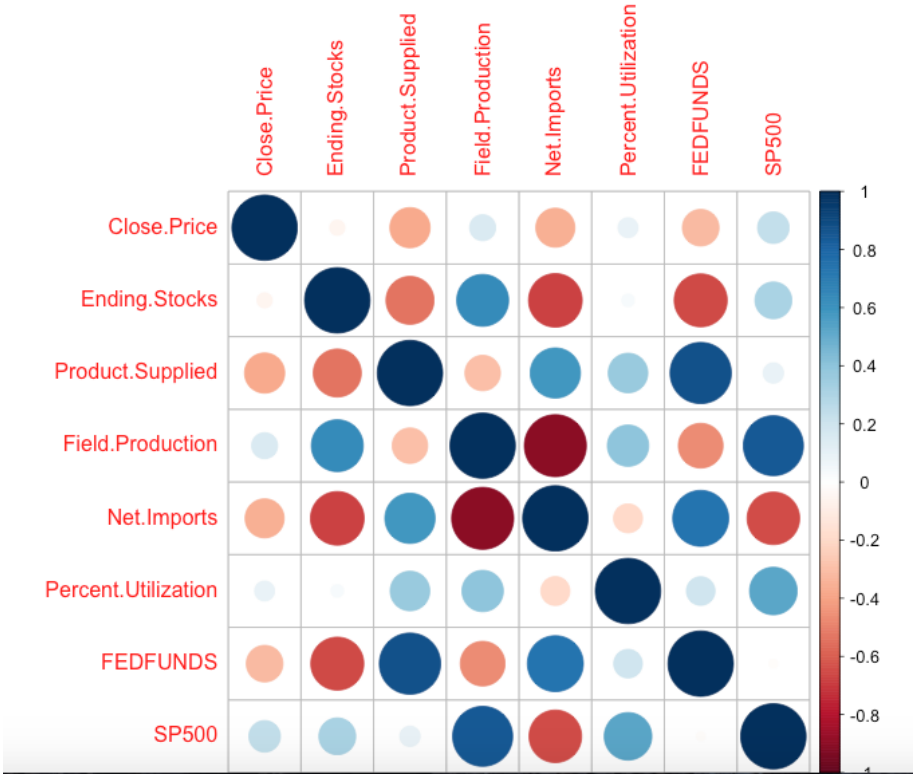
- **Konačne mjesečne zalihe:** zalihe postoje za preveniranje nedostatka ponude u slučaju nepredviđenih događaja. Prethodno smo naveli mnoge primjere tih događaja. Razina ovih zaliha ima jasan utjecaj na cijene sirove nafte: povećanjem zaliha cijena bi trebala biti pod pritiskom te obratno. Podaci su također uzeti sa *EIA* službene stranice.

- **S&P 500 index:** Međudjelovanje svjetskih tržišnih indeksa i nafte proučavano je u mnogim radovima. Primjerice, u svom radu huang et al. (23) razmatraju direktnu povezanost financijskih tržišta i naftnih šokova. S&P indeks je najpraćeniji svjetki indeks dionica. Sadrži 500 kompanija srednje i manje tržišne kapitalizacije koje su registrirane u SAD-u. Podaci su skinuti sa službene stranice *FRED-a (Federal Reserve Bank of St. Louis).*

- **Fed Funds kamatna stopa:** Temeljno je mjerilo kamatnih stopa uz LIBOR u Londonu. Američka centralna banka često manipulira kamatom u svrhu stabiliziranja ekonomije. Jako je praćen pokazatelj od strane svih aktera financijskog tržišta. Podaci su također dobiveni sa službene stranice *FRED-a*.

4.2. Analiza korelacije

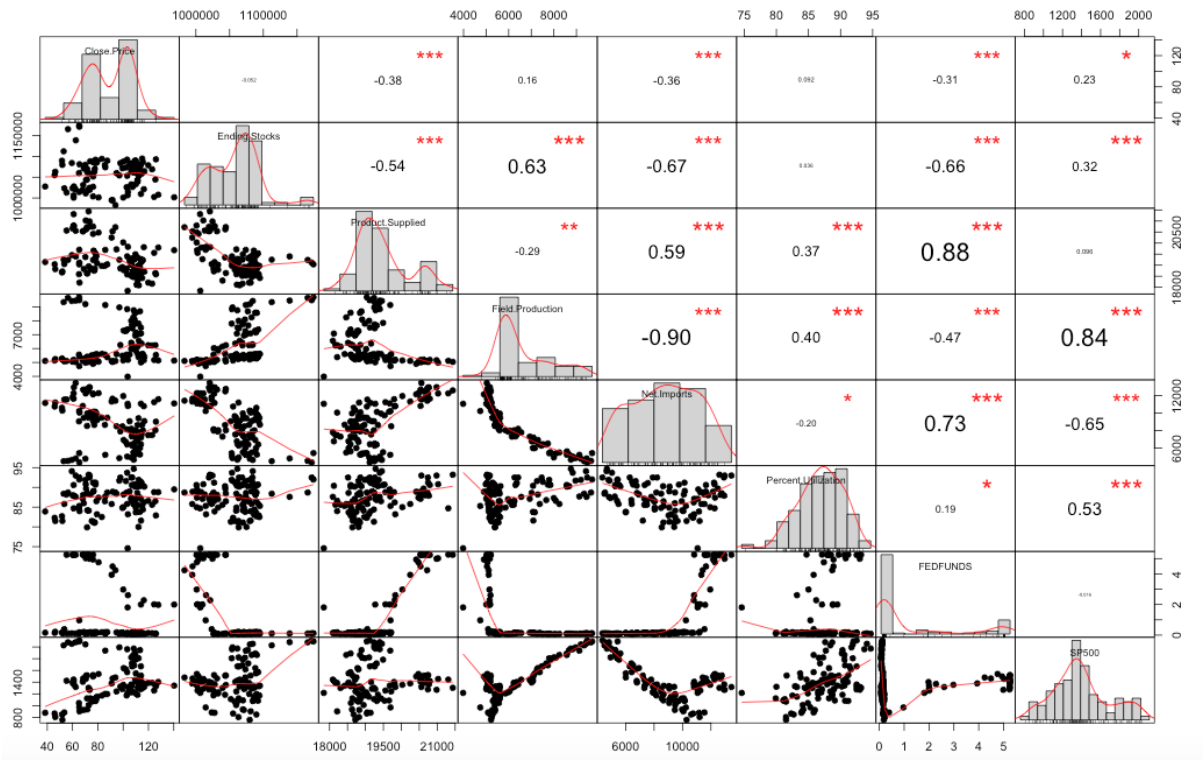
Korelacijskom matricom možemo se bolje upoznati sa setom varijabli te odmah zamjetiti potencijalne kandidate za izbacivanje iz modela. Korelacijskim paketom u statističkom programu R dobiva se korelacijska matrica na slici 17. Odmah se može primjetiti da varijable proizvodnje, potrošnje i neto uvoza iskaču u odnosu na ostale visokom međusobnom korelacijom te korelacijom sa drugim varijablama. Mogu se još zamjetiti snažne korelacije u varijabli zaliha. Iznenađujuće visoku korelaciju ima i varijabla kamatnih stopa.c



Slika 19: Korelacijska matrica

Izvor: Prikaz autora

Znatno detaljniji prikaz možemo dobiti iz *Performance Anallytics* paketa u R-u. Na diagonali nalazi se distribucija svake varijable. Na donjem dijelu diagonale prikaza nalaze se *scatterplot* prikazi sa sa linijom. Na gorenjem dijelu diagonale se nalaze Pearsonovi koeficijenti korelacije sa zvijezdicama koje označavaju razine signifikatnosti. p-vrijednosti redom 0, 0.001, 0.01, 0.05, 0.1 su predstavljene zvijezdicama redom od tri zvijezdice do nula. Vidimo da neto uvoz ima najveći broj statistički značajnih korelacija, što postavlja varijablu kao prvog kandidata za izbacivanje iz modela.



Slika 20: Grafički prikaz distribucija, *scatterplotova*, koeficijenta korelacije i p-vrijednosti

Izvor: Prikaz autora

4.2. Linearna regresija

Korištenjem stepwise metodom (prema naprijed i prema nazad) na temelju Akaikeovog informacijskog kriterija provjerit ćemo koji model najbolje objašnjava našu nezavisnu varijablu. Stepwise algoritam će izbacivati varijable do kse ne postignu sve kombinacije te na temelju kriterija odrediti koji je model optimalan. Rezultat stepwise AIC metode pokazuju da je postojeći model optimalan prema navedenom kriteriju.

```
> step$anova
Stepwise Model Path
Analysis of Deviance Table

Initial Model:
data4$Close.Price ~ data4$Ending.Stocks + data4$Product.Supplied +
    data4$Field.Production + data4$Net.Imports + data4$Percent.Utilization +
    data4$FEDFUNDS + data4$SP500

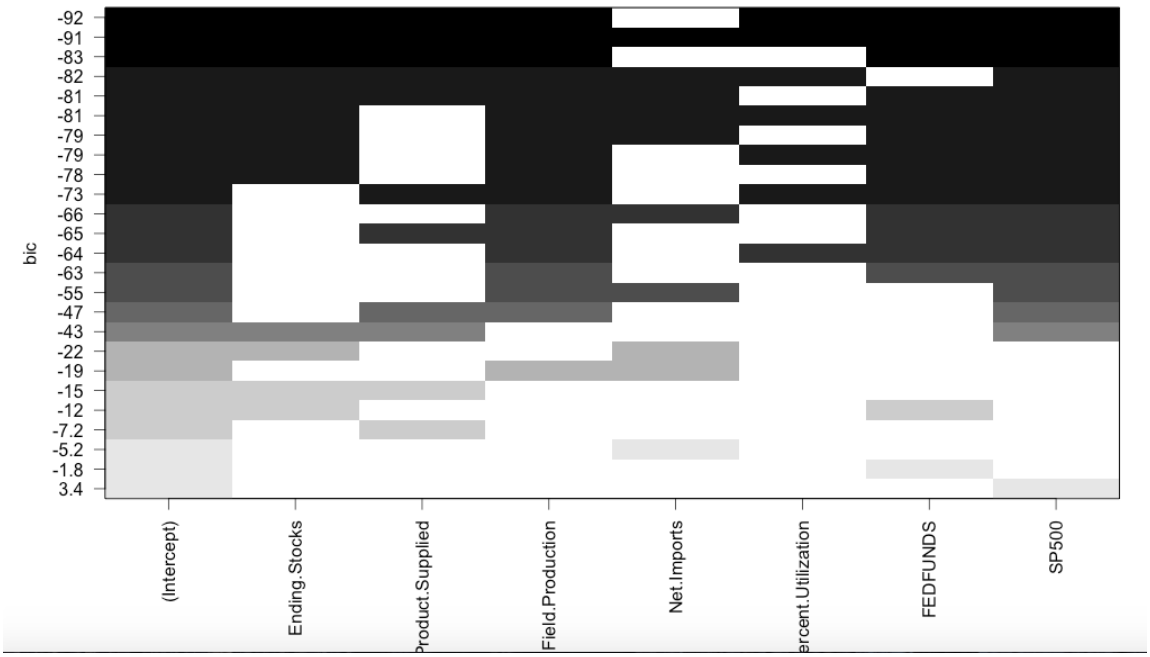
Final Model:
data4$Close.Price ~ data4$Ending.Stocks + data4$Product.Supplied +
    data4$Field.Production + data4$Net.Imports + data4$Percent.Utilization +
    data4$FEDFUNDS + data4$SP500


  Step Df Deviance Resid. Df Resid. Dev      AIC
1                       100   17310.97 564.3121
```

Slika 21: Output *stepwise AIC* metode

Izvor: Prikaz autora

Na sljedećem prikazu vidimo modele sa najnižim Bayesovim informacijskim kriterijom. Prema kriteriju, Neto Uvoz je varijabla koju treba izbaciti što potvrđuje rezultate korelacije. Treba naglasiti da jednake rezultate dobivamo i na temelju R2 kriterija.



Slika 22: Grafički prikaz optimalnih modela prema Bayesovom informacijskom kriteriju

Izvor: Prikaz autora

Prema svemu priloženom, izbacivanje varijable neto uvoza se pokazuje najboljim pristupom. Rezultati našeg konačnog regresijskog modela vidljivi su na sljedećem prikazu. Vidljivo je da svi koeficijenti imaju niske p-vrijednosti. R2 vrijednost je 66,63%.

```
Call:
lm(formula = data4$Close.Price ~ data4$Ending.Stocks + data4$Product.Supplied +
    data4$Field.Production + data4$Percent.Utilization + data4$FEDFUNDS +
    data4$SP500)

Residuals:
    Min      1Q  Median      3Q     Max
-36.466  -6.946  -0.321   7.599  28.848

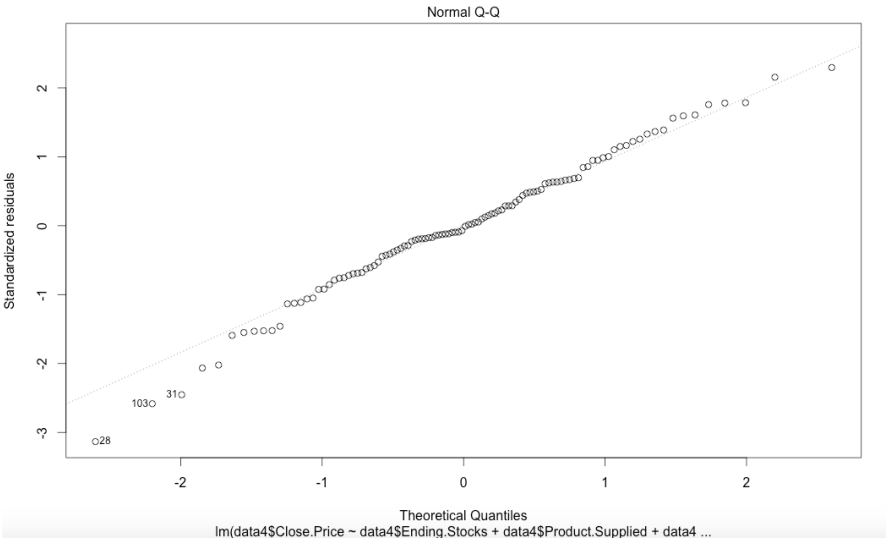
Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)               5.570e+02  8.753e+01   6.363 5.84e-09 ***
data4$Ending.Stocks      -2.709e-04  5.441e-05  -4.978 2.65e-06 ***
data4$Product.Supplied   -1.610e-02  3.745e-03  -4.300 3.94e-05 ***
data4$Field.Production   -2.866e-02  3.434e-03  -8.347 3.82e-13 ***
data4$Percent.Utilization  1.694e+00  4.550e-01   3.723 0.000324 ***
data4$FEDFUNDS           -1.145e+01  1.947e+00  -5.881 5.30e-08 ***
data4$SP500               1.244e-01  1.245e-02   9.994  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 13.31 on 101 degrees of freedom
Multiple R-squared:  0.685,     Adjusted R-squared:  0.6663
F-statistic: 36.61 on 6 and 101 DF,  p-value: < 2.2e-16
```

Slika 23: Grafički prikaz optimalnih modela prema Bayesovom informacijskom kriteriju

Izvor: Prikaz autora

Na Q-Q prikazu reziduala možemo zamjetiti da tri opservacije iskaču: 28, 31 i 103. Prve dvije se nalaze u razdoblju naglog pada u financijskoj krizi 2008. godine dok posljednja je dio značajnog pada 2015. godine zbog fundamentalnih razloga. Ako ignoriramo dva krizna razdoblja u našem vremenskom nizu, gdje sve varijable prezentiraju abnormalno ponašanje, pretpostavka normalnosti reziduala vrijedi.



Slika 24: Q-Q grafički prikaz reziduala

Izvor: Prikaz autora

5. ANALIZA I USPOREDBA REZULTATA

5.1. Sažeti rezultati modela

Na temelju *MSE* (*Mean Squarred Error*) i *MAE* (*Mean Average Error*) usporedit će se prognostički modeli. Pobliže će se objasniti kriteriji odabira modela.

MSE (*Mean Squared Error*) je mjera kvadrirane razlike između (modelom) predviđene vrijednosti te realizirane vrijednosti te potom pronalazi prosječnu središnju vrijednost navedenih razlika. *MSE* predstavlja prosječnu razliku kvadrata reziduala (u kalkulacijama unutar uzorka), odnosno predikcijskih grešaka(u kalkulacijama izvan uzorka). Izraz za *MSE* je sljedeći:

$$MSE = \frac{1}{n}\sum_{j=1}^{n}((\hat{Y}_t(j) - Y_j)^2) \quad (15)$$

MAE (*Mean Absolute Error*) računa središnju vrijednost od apsolutnih vrijednosti reziduala/ predikcijskih grešaka, umjesto kvadrirane vrijednost (koje računa *MSE*). Izraz za *MAE* je sljedeći:

$$MSE = \frac{1}{n}\sum_{j=1}^{n}(|(\hat{Y}_t(j) - Y_j)|) \quad (16)$$

Za usporedbu ARIMA modela odabrani su modeli s najvišim AIC kriterijem. Vidljivo je da nema konzistentnosti kroz vremenske periode ili kriterije ocjenjivanja te je vrlo teško donijeti zaključak iz priloženih rezultata. Također se može primjetiti da optimalni modeli (oni odabrani prema AIC kriteriju u prethodnom poglavlju) imaju *MSE* i *MAE* veličine jako blizu pobjednicima. Sezonski model je odabran prema *MSE* kriteriju u dužem vremenskom nizu, što upućuje na neke oblike sezonskih utjecaja unutar niza, iako je vrlo moguće da su rezultati izobličeni zbog događanja 2008. i 2014. godine.

Tablica 1: ARIMA, vremenski niz 10 godina

10 y	ARIMA(1,1,2)	ARIMA(0,1,0)	ARIMA(1,1,1)	ARIMA(2,1,1)	ARIMA(1,1,2)(0,0,1)
MSE	3.56746	3.605123	3.576464	3.567729	3.564014
MAE	2.605037	2.624842	2.612777	2.604611	2.60741
	ARIMA(1,1,1)(0,0,1)	ARIMA(1,1,1) with drift			
MSE	3.572679	3.576459			
MAE	2.614547	2.61318			

Izvor: Prikaz autora

Tablica 2: ARIMA, vremenski niz 5 godina

5 y	ARIMA(1,1,0)	ARIMA(2,1,0)	ARIMA(0,1,0)	ARIMA(0,1,1) with drift	ARIMA(1,1,0) with drift
MSE	1.50225	1.501536	1.504544	1.502088	1.502232
MAE	1.102543	1.103414	1.105904	1.102437	1.102462

Izvor: Prikaz autora

Usporedba GARCH modela za predviđanje volatilnosti vrši se također višedimenzijski: svi analizirani modeli za petogodišnju vremenski niz dnevnih opservacija te desetogodišnji vremenski niz tjednih opservacija posebno. Također ćemo usporedbu raspodjeliti na vremenske periode predviđanja: 3, 10 i 20 te koristiti *Out of sample* performanse grešaka. Crvenom bojom naznačeni su modeli sa najnižim *MSE i MAE* vrijednostima.

Tablica 3: GARCH vremenski niz 5 godina, broj razdoblja = 20

n=20	sGARCH (1,1)	apARCH (1,1)	gjrGARCH(1,1)	eGARCH(1,1)
MSE	1.1437689	1.1443143	1.1443100	1.144537
MAE	0.7959618	0.7959635	0.7959567	0.795826

Izvor: Prikaz autora

Tablica 4: GARCH vremenski niz 10 godina, broj razdoblja = 20

n=20	sGARCH (1,1)	apARCH (1,1)	gjrGARCH(1,1)	eGARCH(1,1)	sGARCH(1,1)+ ARIMA(1,1,2)
MSE	9.898379	9.853239	9.857671	9.873423	9.710014
MAE	2.615879	2.602511	2.603441	2.603797	2.603480
	apARCH(1,1)+ ARIMA(1,1,2)	gjrGARCH(1,1)+ ARIMA(1,1,2)			
MSE	9.618563	9.616116			
MAE	2.586375	2.586436			

Izvor: Prikaz autora

Tablica 5: GARCH vremenski niz 5 godina, broj razdoblja = 10

n=10	sGARCH (1,1)	apARCH (1,1)	gjrGARCH(1,1)	eGARCH(1,1)
MSE	1.3846394	1.3854871	1.3855582	1.3871095
MAE	0.9038803	0.9038658	0.9038863	0.9041843

Izvor: Prikaz autora

Tablica 6: GARCH vremenski niz 10 godina, broj razdoblja = 10

n=10	sGARCH (1,1)	apARCH (1,1)	gjrGARCH(1,1)	eGARCH(1,1)	sGARCH(1,1)+ ARIMA(1,1,2)
MSE	7.689310	7.605251	7.615853	7.877231	7.910740
MAE	2.228275	2.199954	2.203411	2.219395	2.249183
	apARCH(1,1)+ ARIMA(1,1,2)	gjrGARCH(1,1)+ ARIMA(1,1,2)			
MSE	7.804610	7.910740			
MAE	2.219062	2.249183			

Izvor: Prikaz autora

Tablica 7: GARCH vremenski niz 5 godina, broj razdoblja = 5

n= 3	sGARCH (1,1)	apARCH (1,1)	gjrGARCH(1,1)	eGARCH(1,1)
MSE	1.979310	1.977622	1.977451	1.973307
MAE	1.226888	1.228207	1.228168	1.228175

Izvor: Prikaz autora

Tablica 8: GARCH vremenski niz 10 godina, broj razdoblja = 5

n= 3	sGARCH (1,1)	apARCH (1,1)	gjrGARCH(1,1)	eGARCH(1,1)	sGARCH(1,1)+ ARIMA(1,1,2)
MSE	7.811807	7.583352	7.614847	7.622891	8.018528
MAE	2.220463	2.178350	2.183890	2.189010	2.231270
	apARCH(1,1)+ ARIMA(1,1,2)	**gjrGARCH(1,1)+ ARIMA(1,1,2)**			
MSE	7.784021	7.805209			
MAE	2.189986	2.192448			

Izvor: Prikaz autora

Ako vremenske nizove promatramo odvojeno, za petogodišnji niz sa dnevnim vremenskim opservacijama najboljim se pokazao sGARCH model (baziran na AR(1) diferenciranog niza). Model je bio optimalan u svim razdobljima predviđanja: 5,10 i 20, što potvrđuje Cheongove (2009) zaključke.

Rezultati desetogodišnjeg niza sa tjednim opservacijama se razlikuju, apARCH je jasno dominantan model te u većini slučajeva dovoljno je model bazirati na običnom AR(1) procesu diferenciranog niza. Intuitivno objašnjenje može se pronaći u asimetričnim volatilnostima koje apARACH uključuje pri modeliranju. Naime, u desetogodišnjem vremenskom nizu, osim snažnog pada cijena u 2014. godini (također dio drugog niza), uključeno je i iznimno volatilno razdoblje financijske krize 2008. godine. apARACH potvrđuje rezultate Mohammadi i Sua (2010). Stoga možemo zaključiti da je apARCH model dominantan ako vremenski niz sadrži veći broj geopolitičkih utjecaja, odnosno ako se odnosi na duže vremensko razdoblje, dok jednostavan GARCH model je dovoljan u ostalim slučajevima. Ako se vratimo na petogodišnji vremenski niz, možemo vidjeti da apARCH

model ima jako niske kriterije ocjenjivanja te je i u jednom slučaju imao i najniži *MAE*. Ako treba odabrati pobjednika, onda je to definitivno apARCH.

Dodatno, možemo zaključiti da modeliranjem ARCH modela prema prethodno definiranim optimalnim ARIMA modelima nije pružilo nikakvu dodatnu vrijednost. Stoga zaključujemo da je ARCH modele najbolje modelirati prema najjednostavnijem ARIMA(1,0,0) procesu.

6. ZAKLJUČAK

Ovim radom nastojalo se modelirati ponašanje vremenskog niza cijene sirove nafte, jednog od najvolatilnijih i iznimno teško predvidivih proizvoda financijske industrije. Kako sirovu naftu karakteriziraju snažni geopolitički i fundamentalni utjecaji, uključen je veliki broj kriterija i modela. Različitim varijantama ARIMA modela uključeni su utjecaji lagiranih varijabli te pomičnih prosjeka. Linearnom regresijom modelirali su se utjecaji eksternih varijabli. Konačno, GARCH modelima uveden je sloj volatilnosti i varijance.

Samom činjenicom da su ARIMA modeli pokazali nekonzistentne rezultate te da se nijedan model nije pokazao dominantim u konkretnim scenarijima, možemo zaključiti da jednostavnom diferencijom i primjenom AR(1) modela dobivamo rezultate koji mogu parirati kompleksnijim ARIMA oblicima. Također dobivamo jednostavan input za GARCH modele.

Zbog nedostatka podataka, linearna regresija nije mogla biti direktno uspoređena sa drugim modelima, već se promatrala na mjesečnoj razini. Vidjeli smo da odabirom fundamentalno povezanih nezavisnih varijabli, regresija također ima potencijal preciznog opisivanja zavisne varijable. Na temelju kriterija kao *AIC*, *R2* i *BIC* došli smo do najboljeg modela linearne regresije.

Konačno, pri modeliranju volatilnosti dobiveni su jako zanimljivi rezultati. apARCH model dao je najbolje rezultate prema *MSE* i *MAE* kriterijima. Što je veća asimetrična volatilnost niza, to je apARCH model precizniji u odnosu na ostale. Upravo zato u vremenskom nizu koji uključuje financijsku krizu, apARCH se pokazao optimalnim. Akteri u proizvodno-potrošnom lancu sirove nafte mogu kratkoročnu volatilnost modelirati prema ovom modelu znajući da je najbolja alternativa te na taj način osigurati svoj rizik u iznimno volatilnom i opasnom tržištu.

POPIS TABLICA

Tablica 1: ARIMA, vremenski niz 10 godina.....44
Tablica 2: ARIMA, vremenski niz 5 godina.....44
Tablica 3: GARCH vremenski niz 5 godina, broj razdoblja = 20 44
Tablica 4: GARCH vremenski niz 10 godina, broj razdoblja = 20 45
Tablica 5: GARCH vremenski niz 5 godina, broj razdoblja = 10 45
Tablica 6: GARCH vremenski niz 10 godina, broj razdoblja = 10 45
Tablica 7: GARCH vremenski niz 5 godina, broj razdoblja = 5 46
Tablica 8: GARCH vremenski niz 10 godina, broj razdoblja = 5 46

POPIS SLIKA

Slika 1: Futures WTI CL, 2012 – 2016. Interactive brokers, NYMEX podaci 12
Slika 2: Deskriptivna statistika za oba vremenska niza.....17
Slika 3: Linijski graf - dnevne opservacije cijena zatvaranja Brent futures ugovora, 31.5.2006. - 25.5.2010. godine 18
Slika 4: Linijski graf - tjedneopservacije cijena zatvaranja Brent futures ugovora, 2.6.2005. - 22.5.2015.....19
Slika 5: Output Augmented Dickey-Fuller testa u statistčkom alatu R.....20
Slika 6: Periodogram – differencirani niz mjesečnih opservacija; 31.5.2006. do 25.5.2010. .21
Slika 7: Periodgram – diferencirani niz mjesečnih opservacija; 2.6.2005. do 22.5.2015.....22
Slika 8: Zivot-Andrews test jediničnog korjena za 10 godišnji vremenski niz (lijevo) i 5 godišnji vremenski niz (desno).....23
Slika 9: AR(1) diferenciranih vremenskih nizova: niz 5 godina (lijevo), niz 10 godina (desno) 24
Slika 10: Grafički prikaz AR(1) prognoze vremenskih nizova: 91 za 5 godišnji niz (lijevo), 52 za 10 godišnji (desno).....25
Slika 11: Ljung-Box test autokorelacije reziduala AR(1) modela.....26
Slika 12: Rezultati filtriranja ARIMA modela na temelju informacijskih kriterija: vremenski niz sa dnevnim opservacijama (lijevo), vremenski niz sa tjednim opservacija (desno) 28

Slika 13: Prognozirane vrijednosti vremenskih nizova: 1 kvartal (91 opservacija) za vremenski niz s dnevnim opservacijama (lijevo), 1 godina (52 opservacije) za vremenski niz sa tjednim opservacija 29
Slika 14: Autokorelacijske funkcije reziduala ARIMA (1,1,0) modela (lijevo) te ARIMA (1,1,2) modela (desno) 29
Slika 15: Ljung-Box test autokorelacije reziduala ARIMA(1,1,0) modela (gore) te ARIMA(1,1,2) modela (dolje) 30
Slika 16: Statistički pokazatelji za sGARCH(1,1) model primjenjen na ARIMA(1,1,0) za petogodišnju vremensku seriju dnevnih opservacija, prvi dio. 32
Slika 17: Statistički pokazatelji za sGARCH(1,1) model primjenjen na ARIMA(1,1,0) za petogodišnju vremensku seriju dnevnih opservacija, drugi dio. 33
Slika 18: Projecirane vrijednosti vremenske serije i standardne devijacije ARIMA(1,1,0)-GARCH(1,1) modela 34
Slika 19: Korelacijska matrica 38
Slika 20: Grafički prikaz distribucija, *scatterplotova*, koeficijenta korelacije i p-vrijednosti 39
Slika 21: Output *stepwise AIC* metode 40
Slika 22: Grafički prikaz optimalnih modela prema Bayesovom informacijskom kriteriju 40
Slika 23: Grafički prikaz optimalnih modela prema Bayesovom informacijskom kriteriju 41
Slika 24: Q-Q grafički prikaz reziduala 42

LITERATURA:

1. Ahmed R. A., Shabri A. B. (2014): Daily Crude Oil Price Forecasting Model Using ARIMA, Generalized Autoregressive Conditional Heteroscedastic and Support Vector Machines. American Journal of Applied Sciences 11 (3): 425-432
2. Akpanta A. C., Okorie I. E. (2014): Application of Box-Jenkins Techniques in Modelling and Forecasting Nigeria Crude Oil Prices. International Journal of Statistics and Applications. 4 (6), str. 283-291
3. Alquist R. Kilian L. (2007): What do we Learn from the Price of Crude Oil Futures? *University of Michingan and CEPR, Department of Economics.*
4. Alquist R., Kilian L., Vigfusson R. J., (2011). Forecasting the Price of Oil, *International Finance Discussion Papers*, n. 1022
5. Basher S.A., Sadorsky P. (2006). Oil price risk and emerging stock markets, *Global Finance Journal* 17. 224-251
6. Bollerslev T. (1986): Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307-327
7. Deng A., King K., Metz D. (2012): An Econometric Analysis of Oil Price Movements: The Role of Political Events and Economic News, Financial Trading, and Market Fundamentals. Bates White Economic Consulting
8. Economic Research, Federal Reserve bank of St. Louis. [Internet], raspoloživo na: https://fred.stlouisfed.org
9. Federer J. P. (1996). Oil Price Volatility and the Macroeconomy, *Journal of Macroeconomics*, 18(1), pp. 1-26.
10. Fleming, J., Ostdiek, B. (1999). The impact of energy derivatives on the crude oil market. *Energy Economics*, 21, 135-167
11. Ghalanos A. (2015): Introduction to the rugarch package (Version 1.3-1). [Internet], raspoloživo na: https://cran.r-project.org/web/packages/rugarch/vignettes/Introduction_to_the_rugarch_package.pdf
12. Gupta R., Lux T., Segnon M. (2015): Modeling and Forecasting Crude Oil Price Volatility: Evidence from Historical and Recent Data. University of Pretoria, Department of Economics Working Paper Series.
13. Hamilton, J.D. (2009), "Causes and Consequences of the Oil Shock of 2007-08," Brookings Papers on Economic Activity, 1, Spring, 215-261.

14. Hamilton J. D. (1983). Oil and the Macroeconomy Since World War II, *Journal of Political Economy*, 91, 228-248.
15. He L., Yao T., Zhang Y. (2015): Forecasting crude oil market volatility: can the Regime Switching GARCH model beat the single-regime GARCH models? [Internet], raspoloživo na: http://arxiv.org/abs/1512.01676
16. Huang R. D., Masulis R. W., Stoll H. R. (1996): Energy shocks and financial markets. Journal of Futures Markets, 16, 3956.
17. Hyndiman R.J., Khandakar Y. (2008): Automatic Time Series Forecasting: The forecast Package for R. Journal of Statistical Software. Volume 27 (3)
18. Jarque, C. M., Bera, A.K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, 6(3), 255259.
19. Kaufmann, R.K., Dees, S., Karadeloglou, P., Sanchez, M. (2004). Does OPEC matter? An econometric analysis of oil prices. *The Energy Journal*, 25(4), 67-90.
20. Lam. D. (2013). Time Series Modelling of Monthly WTI Crude Oil Return, *Thesis subimitted for the degree of Master of Science in Mathematical and Computational Finance.*
21. Lutkepohl H., Xu F., (2009): The Role of Log Transformation in Forecasting Economic Variables
22. Manescu C., Robays I. V.(2014). Forecasting the Brent Oil Price. Addressing Time-Variation in Forecast Performance. *Working Paper Series, ECB.* No. 1735
23. Maximo C. (2011): Markov-switching models and the unit root hypothesis in real US GDP. Economic Letters, Volume 112 (2), str. 161-164
24. Narayan, P.K., Narayan S. (2007). Modelling oil price volatility. *Energy Policy*, 35, 6549-6553.
25. Park W. J., Ratti. R.A.(2008). Oil price shocks and Stock markets in the U.S. and 13 European Countries, *Energy Economics*, 30, 2587-2608.
26. Perron P. (1988): The Great Crash, the Oil Price Shock and the Unit Root Hypothesis. Econometric Research Program Research Memorandum, 338
27. Pindyck R. S. (1999). The Long-Run Evolution of Energy Prices, *Center of Energy and Environmental Policy Research*, 99-001. WP
28. Pirog R. (2005) World Oil Demand and its Effect on Oil Prices. CRS Report for Congress, order code RL32530

29. Salisu A. A., Fasanya I.O. (2012) Comparative Performance of Volatility Models for Oil Price. *International Journal of Energy Economics and Policy*, Vol 2.,No. 3, pp.167-183.
30. Shih M. L. , Shyh W. C. (2014): Non-linear Dynamics in international resource markets: Evidence from regime switching approach. Research in International Business and Finance, 30 (C), str. 233-247
31. U.S. Energy Information Administration. [Internet], raspoloživo na: http://www.eia.gov
32. Zivot E. (2003): Lectures on Structural Change. Department of Economics, University of Washington. [Internet], raspoloživo na: http://faculty.washington.edu/ezivot/book/structuralchangeslides1.pdf

SAŽETAK

Ovaj rad bavi se problematikom modeliranja cijene i volatilnosti sirove nafte na dva seta podataka: desetogodišnji vremenski niz s tjednim opservacija te petogodišnji vremenski niz s dnevnim opservacijama. Započinje se analizom fundamentalnih karakteristika proizvoda i tehnikama deskriptivne statistike. Nakon pobližeg proučavanja dinamike vremenskih nizova, testiranjem stacionarnosti, sezonalnosti i strukturnih lomova, potrebnim transformacijama podaci se pripremaju za modeliranje. Na temelju odabranih informacijskih kriterija biraju se optimalni ARIMA modeli čije su karakteristike detaljnije proučene. Biranjem nezavisnih varijabli i testiranjem korelacije, gradi se optimalni model linearne regresije koji se potom dodatno razrađuje. Kao nastavak na ARIMA proces, modelira se osnovni GARCH model te se vrši naknadna procjena modela. Modeliraju se i različite varijante GARCH-a: apARCH, eGARCH i gjrGARCH.

Konačno, različiti ARIMA modeli se međusobno uspoređuju na temelju MSE i MEA kriterija. Rezultati ne pokazuju prednost u odabiru konkretnog modela, čime se naglasak prebacuje na odabir jednostavnijih modela, kao što je AR(1). GARCH modeli se također uspoređuju na temelju istih kriterija. Rezultati pokazuju znatnu prednost u korištenju apARCH modela s naglaskom na desetogodišnji vremenski niz. Regularni sGARCH predstavlja drugu najbolju alternativu. GARCH modeli pokazuju bolje rezultate u modeliranju vremenskih nizova sirove nafte u odnosu na ARIMA modele zbog volatilne prirode proizvoda.

SUMMARY

This paper examines the problematic of modeling crude oil price and volatility on two sets of data: a ten year time series with weekly observations and a five year series with daily observations. It begins by analyzing the fundamental characteristics of the product and by employing methods of descriptive statistics. After closely studying the dynamics of the series by testing for stationarity, seasonality and structural breaks, with the necessary transformations the data is prepared for modeling. Based on the selected information criterion, the optimal ARIMA model is chosen, from whoms characteristics are studied further. By choosing independent variables and testing for correlation, an optimal linear regression model is built which is then furtherly analyzed. As an addition to the ARIMA process, a basic GARCH model is built, followed by postestimation. Different variants of GARCH are also modeled: apARCH, eGARCH and gjrGARCH.

Finally, different ARIMA models are compared based on the MSE and MEA criteria. The results don't show an advantage in choosing a concrete model, which therefore stresses the importance of choosing a simpler model, like AR(1). GARCH models are also compared based on the same criterion. The results show an advantage in using the apARCH model with emphasis on the ten year time series. The regular sGARCH is the second best alternative model choice. GARCH models show better results in modeling crude oil time series in comparison to ARIMA models because of the high volatility nature of the product.