Univerzitet u Beogradu
Matematički fakultet u Beogradu

Prognoziranje vremenskih serija – Modeli i izbor modela

-master rad-

Mentor:
dr Jelena Jocković

Kandidat:
Petar Korović (1009/2015)

Beograd,
septembar 2016. godine

Sadržaj

1 Uvod

Pod vremenskom serijom podrazumevamo uređeni niz opservacija. Uređivanje se najčešće vrši u odnosu na vreme u jednakim vremenskim intervalima. Serije koje će biti razmatrane u ovom radu će imati ovakvu karakterizaciju. Posle formalnog uvođenja pojma slučajnog procesa vremenska serija se definiše kao jedna realizacija pomenutog procesa. Vremenske serije srećemo u različitim sferama ljudskog delovanja. Na primer u demografiji je to, između ostalog stopa prirodnog priraštaja, u ekonomiji fluktuacije deviznog kursa, mesečno kretanje industrijske proizvodnje, uvoza, izvoza i cena; kvartalne podatke o ugostiteljstu i godišnju vrednost društvenog proizvoda,…

U radu pored glavnog cilja prognoziranja, predstavljanja i opisa prognostičkih modela, kao i praktičnog dela takođe će biti reči i o analizi vremenskih serija, slučajnim procesima i linearnim procesima, koji su neophodni radi uvođenja u problematiku prognoziranja.

Prognoziranje vremenskih serija odnosno određivanje budućeg toka je jedan od jako važnih ako ne i najvažniji cilj analize vremenskih serija. Na primer, na nivou privrede kao celine od interesa je prognozirati kretanje društvenog proizvoda, jer se na osnovu dobijene prognoze procenjuje obim i udeo sredstava za zajedničke i opšte potrebe. Takođe možemo zaključiti da formirana prognoza društvenog proizvoda povratno deluje na nosioca ekonomske politike. Dakle, direktno proizilazi da pogrešna prognoza za sobom povlači i druge neželjene efekte, kao što su materijalne posledice u vidu dodatnih troškova prouzrokovanih pogrešno donetim odlukama.

Posle teorijskog dela i predstavljanja modela uslediće i praktičan deo rada, prognoziranja vremenskih serija: *industrijske proizvodnje*, *potrošačkih cena*, *prometa (trgovina) robe u stalnim cenama*, *izvoza robe u evrima* i *uvoza robe u evrima*. Kako je većina naših vremenskih serija nestacionarna model koji ih najbolje opisuje jeste ARIMA model (koji diferenciranjem seriju svodi na stacionarnu) koji kombinovan sa aritmetičkom sredinom i ekspertskom korekcijom daje najbolji “recept” za kvalitetnu prognozu, a više o tome će biti reči u glavama 4 i 5.

2 Analiza vremenskih serija

Vremenske serije imaju široku primenu u praksi, koriste se u ekonometriji, statistici, finansijama, kao i u raznim svakodnevnim situacijama, pa zapravo nismo ni svesni koliko ih koristimo (npr. u praćenju prometa - pazara neke radnje, vremenskoj prognozi, itd), tj. u svakoj oblasti koja uključuje merenje neke pojave kroz vreme. Glavni ciljevi u analizi vremenskih serija su *opisivanje*, *objašnjavanje* i *predviđanje – prognozioranje* vremenske serije. Posmatranjem grafički predstavljene vremenske serije mogu se uočiti određene determinističke i slučajne komponente. Jedna od važnih komponenti je *trend*, koji izražava dugoročnu tendenciju razvoja serije. U zavisnosti od toga da li serija raste ili opada, trend može biti rastući ili opadajući. Pored ovoga, trend takođe može biti deterministički ili stohastički, što zavisi od toga da li se kretanje vremenske serije može predvideti na osnovu prethodnih vrednosti ili ne. Pored trenda, kao deterministička komponenta često se javlja i *sezonska komponenta*, ona je rezultat dejstva faktora koji se javljaju u određenim periodima, sezonama. Vremenska serija sadrži i slučajnu komponentu, koja se naziva *beli šum*. Veza ovih komponenata može biti *aditivna*, *multiplikativna* ili *mešovita* (kombinacija aditivne i multiplikativne). Pored pomenutih, kao komponente vremenske serije često se javljaju i: *kalendarska komponenta* (uticaj kalendara), *iregularna komponenta* (događaji koji se ne mogu predvideti, retko se dešavaju, a kad se dese značajno utiču na seriju npr. zemljotresi, ratovi, poplave, suša, političke mere itd).

Analiza vremenskih serija predstavlja statističku disciplinu u kojoj se zaključivanje bitno razlikuje od uobičajnog načina statističkog zaključivanja. Osnovni pojam kod statističkog zaključivanja je *prost slučajan uzorak*, pod kojim smatramo skup od n nezavisnih i jednako raspodeljenih slučajnih promenljivih, dok kod vremenskih serija se, takođe, razmatra skup slučajnih promenljivih za koje se pretpostavlja da su međusobno zavisne i veoma često i korelisane. Ovu zavisnost koristimo radi formiranja modela vremenske serije, a zatim model koristimo da na osnovu prošlih vrednosti formiramo - ***prognoziramo*** buduće.

Analiza standardnih vremenskih serija počiva na specifičnim osobinama vremenskih serija, kao što su *stacionarnost*, *autokorelisanost*, *beli šum*. Te kvalifikacije serije govore o tome da li je stacionarna, kako izgleda njena *autokorelaciona* i *autokovarijaciona* funkcija, kao i koju raspodelu ima beli šum. Jedan od važnih modela je ***autoregresioni model pokretnih proseka, ARMA***. U ovom radu biće objašnjeni navedeni pojmovi kao i njihova primena. Pored ARMA modela biće predstavljeni i drugi relevantni modeli. Ukoliko vremenska serija nije stacionarna, potrebno je prvo svesti na stacionarnu, a zatim je modelirati, najčešće, ARMA modelom. Model koji je adekvatan da pokrije oba slučaja (stacionarne i nestacionarne serije) jeste **ARIMA** model.

2.1 Slučajni procesi

Ovde ćemo izložiti neke osnovne definicije vezane za stacionarne procese i njihove osobine ([1]).

Označimo sa Ω prostor ishoda nekog eksperimenta. Podskupove Ω nazivamo događaji, a klasu svih događaja označavamo sa $\mathcal{A}$. Klasi događaja pripada skup Ω (siguran događaj), prazan skup $\varnothing$ i klasa je zatvorena u odnosu na primenu skupovnih operacija.

Definicija 1 *Klasa* $\mathcal{A}$ *podskupova prostora ishoda* Ω*, koja ima sledeća svojstva:*

1. $\Omega \in \mathcal{A}$,
2. Ako $A \in \mathcal{A}$, onda $\overline{A} \in \mathcal{A}$,
3. Ako $A_1 \in \mathcal{A}, A_2 \in \mathcal{A}, A_3 \in \mathcal{A},...,$ onda $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$,

zove se σ-algebra događaja. Elementi σ-algebre zovu se slučajni događaji ili samo događaji.

Uređeni par $(\Omega, \mathcal{A})$ zove se merljiv prostor.

Definicija 2 *Funkcija* $P : \mathcal{A} \to \mathbf{R}$ *je verovatnoća na merljivom prostoru* $(\Omega, \mathcal{A})$*, ako ima sledeća svojstva:*

1. $P(\Omega) = 1$,
2. $P(A) \geq 0$ za svaki događaj $A \in \mathcal{A}$,
3. Ako su $A_1, A_2, A_3,...$ događaji iz $\mathcal{A}$, takvi da za različite i i j važi $A_i \bigcap A_j = \varnothing$,

onda važi jednakost

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n).$$

Definicija 3 *Uređena trojka* $(\Omega, \mathcal{A}, P)$*, gde je*

- Ω - prostor ishoda slučajnog eksperimenta,
- $\mathcal{A}$ - σ-algebra podskupova skupa Ω,
- P - verovatnoća definisana na σ-algebri $\mathcal{A}$,

zove se prostor verovatnoća ili verovatnosni model razmatranog slučajnog eksperimenta.

Definicija 4 *Neka je dat prostor verovatnoća* $(\Omega, \mathcal{A}, P)$*, i neka je* $(\mathbf{R}, \mathcal{B})$ *uređeni par čije su komponente skup realnih brojeva i Borelova* σ*-algebra podskupova skupa realnih brojeva. Funkcija* $X : \Omega \to \mathbf{R}$ *zove se slučajna veličina, ako je merljiva u odnosu na* σ*-algebre* $\mathcal{A}$ *i* $\mathcal{B}$*, tj. ako za svaki Borelov skup* $B \in \mathcal{B}$ *važi*

$$X^{-1}(B) = \{\omega : X(\omega) \in B\} \in \mathcal{A}.$$

Definicija 5 *Familija slučajnih elemenata* $\{X_t, t \in T\}$ *definisanih na nekom prostoru verovatnoća* $(\Omega, \mathcal{A}, P)$*, gde je T beskonačan skup, zove se slučajna funkcija.*

Skup T iz prethodne definicije zove se parametarski skup. Ako je $T = \mathbf{R} = (-\infty,+\infty)$, $T = [0,+\infty)$ ili $T = [a,b] \subset \mathbf{R}$, onda se parametar $t \in T$ najčešće interpretira kao vreme, a slučajna funkcija $\{X_t, t \in T\}$ zove se ***slučajan proces*** sa neprekidnim vremenom. Ako je $T \subset \mathbf{Z}$, gde je $\mathbf{Z}$ skup celih brojeva, onda se slučajna funkcija $\{X_t, t \in T\}$ zove slučajan proces sa diskretnim vremenom.

2.2 Stacionarne vremenske serije

Definicija 6 *Vremenska serija X_t je jedna realizacija realnog slučajnog procesa $\{X_t, t \in T\}$ sa neprekidnim ili diskretim parametrom pri čemu slučajne veličine X_t i X_s, $t \neq s$ nisu nekorelisane.*

Ukoliko se svojstva vremenske serije tokom vremena ne menjaju onda je data serija stacionarna. U suprotnom ona je nestacionarna. Kod vremenskih serija stacionarnost igra jednu od veoma bitnih uloga. U zavisnosti od toga da li je vremenska serija stacionarna ili ne, biramo različite statističke metode za analizu, pa tako vremenske serije delimo na stacionarne i nestacionarne ([2], [4]).

Navodimo dve osnovne definicije stacionarnosti.

Definicija 7 *(Stroga stacionarnost) Vremenska serija X_t je strogo stacionarna ako za bilo koja dva prirodna broja n i k i bilo koju n-torku prirodnih brojeva $(t_1,...,t_n)$ slučajni nizovi $(X_{t_1},...,X_{t_n})$ i $(X_{t_1+k},...,X_{t_n+k})$ imaju istu funkciju raspodele.*

Dakle, strogo stacionarna vremenska serija je ona vremenska serija čija se svojstva ne menjaju prilikom translacije u vremenu.

Definicija 8 *(Slaba stacionarnost) Za vremensku seriju X_t kažemo da je slabo stacionarna ukoliko važe sledeći uslovi:*

1. $EX_t^2 < \infty$, $\forall t \in Z$,
2. $EX_t = \mu$, $\forall t \in Z$,
3. $cov(X_t, X_{t-k}) = E(X_t - \mu)E(X_{t-k} - \mu) = \gamma(k)$, $\forall t,k \in Z$,

gde je $\gamma(\cdot)$ autokovarijaciona funkcija.

Obe definicije nam govore da je ponašanje vremenske serije slično u svakom posmatranom vremenskom trenutku. Bilo koja promena *očekivanja*, *disperzije* ili *kovarijacije* u posmatranim trenucima je u suprotnosti sa stacionarnošću.

Lako se može zaključiti da iz stroge stacionarnosti sa konačnom disperzijom sledi slaba stacionarnost. Međutim, kada proces ima beskonačnu disperziju, on može biti strogo stacionaran, ali ne važi slaba stacionarnost.

2.3 Autokovarijaciona i autokorelaciona funkcija

Jedna od osnovnih stavki u analizi vremenskih serija je ispitivanje korelacije između slučajne promenljive i njenih prošlih vrednosti. Za ispitivanje korelacije koristimo autokorelacionu i autokovarijacionu funkciju ([3]).

Definicija 9 *(autokovarijaciona funkcija) Autokovarijaciona funkcija* $\gamma(h)$ *stacionarnog procesa* $(X_t)_{t\in\mathbf{Z}}$ *definisana je sledećom formulom*

$$\begin{aligned}\gamma(h) &= Cov(X_t, X_{t-h})\\ &= E(X_t - E(X_t))(X_{t-h} - E(X_{t-h}))\\ &= E(X_t X_{t-h} - X_t E(X_{t-h}) - X_{t-h}E(X_t) + E(X_t)E(X_{t-h}))\\ &= E(X_t X_{t-h}) - E(X_t)E(X_{t-h}), \quad \forall h \in \mathbf{Z}.\end{aligned}$$

Stacionarnost je iskorištena u poslednjoj jednakosti. Potiranje članova je moguće na osnovu stacionarnosti.

Autokovarijaciona funkcija se takođe može definisati kao niz autokovarijacionih koeficijenata na rastojanju *k*. Autokovarijacioni koeficijenti su dati sa $\gamma_k = E(X_t X_{t-k}) - E(X_t)E(X_{t-k})$ za konkretno *k*, a kad *k* prođe skupom **Z** dobijamo autokovarijacionu funkciju.

Definicija 10 *(autokorelaciona funkcija) Autokorelaciona funkcija* $\rho(h)$ *slabo stacionarnog procesa* $(X_t)_{t\in\mathbf{Z}}$ *definisana je sledećom formulom*

$$\rho(h) = \frac{Cov(X_t, X_{t-h})}{\sqrt{D(X_t)}\sqrt{D(X_{X_{t-h}})}} = \frac{Cov(X_t, X_{t-h})}{\sqrt{D(X_t)}} = \gamma(h)/\gamma(0), \quad \forall h \in \mathbf{Z}.$$

Stacionarnost je iskorištena u sređivanjem imenioca, jer predpostavka je da je disperyija konstantna.

Autokorelaciona funkcija se takođe može definisati kao niz autokorelacionih koeficijenata na rastojanju *k*. Autokorelacioni koeficijenti na rastojanju *k* dati su sa $\rho_k = \dfrac{\gamma_k}{\gamma_0}$ za konkretno *k*, a kad *k* prođe skupom **Z** dobijamo autokrelacionu funkciju.

OSOBINE AUTOKOVARIJACIONE I AUTOKORELACIONE FUNKCIJE

1. $\gamma(0) = D(X_t)$, $\rho(0) = 1$.
2. Osobina simetričnosti

$$\gamma(h) = \gamma(-h), \quad \rho(h) = \rho(-h).$$

Dokaz: Na osnovu uslova stacionarnosti sledi

$$\gamma(h) = Cov(X_t, X_{t-h}) = Cov(X_{t+h}, X_t) = Cov(X_t, X_{t+h}) = \gamma(-h).$$

Stacionarnost je iskorištena u drugoj jednkosti (činjenica da se ne menja kroz vreme). Slično dobijamo i za $\rho(h) = \rho(-h)$.

3. $|\gamma(h)| \leq \gamma(0)$ i $|\rho(h)| \leq 1$.

Dokaz: Slučajna promenljiva $Y = \lambda_1 X_t + \lambda_2 X_{t-k}$ ima nenegativnu disperziju, gde su λ_1 i λ_2 proizvoljne konstante. Disperzija slučajne veličine Y je

$$D(Y) = D(\lambda_1 X_t + \lambda_2 X_{t-k}) = (\lambda_1^2 + \lambda_2^2) D(X_t) + 2\lambda_1\lambda_2\gamma(h)$$

za $\lambda_1 = \lambda_2 = 1$ dobijamo $\gamma(h) \geq -D(X_t)$, pa je $\rho(h) \geq -1$,

za $\lambda_1 = 1$, $\lambda_2 = -1$ dobijamo $\gamma(h) \leq D(X_t)$, pa je $\rho(h) \leq 1$.

2.4 Beli šum

Kao što je već rečeno, vremenske serije se sastoje od determinističkih komponenti i slučajne komponente poznatije kao ***beli šum***. Beli šum je slabo stacionaran proces. On je jako bitan, jer su mnogo kompleksniji stacionarni procesi izgrađeni na osnovu njega. Još jedna važna osobina belog šuma je nekorelisanost ([2], [4]).

Definicija 11 *(slab beli šum) Proces* ε_t *se naziva slabi beli šum ako za pozitivnu konstantu* σ^2 *važi sledeće:*

1. $E(\varepsilon_t) = 0, \quad \forall t \in \mathbf{Z}$,
2. $E(\varepsilon_t^2) = \sigma^2, \quad \forall t \in \mathbf{Z}$,
3. $\gamma(h) = cov(\varepsilon_t, \varepsilon_{t+h}) = 0, \quad \forall t, h \in \mathbf{Z}, \quad h \neq 0$.

U prethodnoj definiciji nije data pretpostavka o nezavisnosti. Veoma bitna stavka za vremenske serije je zapravo nekorelisanost procesa beli šum, pa zbog toga je neophodno treću stavku u definiciji zameniti jačom pretpostavkom:

3^*. ε_t i ε_{t+h} nezavisne i jednako raspodeljene.

Definicija 12 *(jaki beli šum) Proces* ε_t *se naziva jaki beli šum ukoliko su promenljive* ε_t *i* ε_{t+h} *nezavisne i jednako raspodeljene, sa očekivanjem 0 i disperzijom* $D(\varepsilon_t) = \sigma^2 = const$.

U praksi se često pretpostavlja da ε_t ima standardizovanu normalnu raspodelu ili Studentovu t-raspodelu.

Beli šum je bitan u analizi vremenskih serija, zbog modeliranja vremenskih serija, kao i zbog toga što on služi za konstruisanje drugih procesa. Pokazatelji belog šuma kao što su autokovarijaciona i autokorelaciona funkcija, odnosno, statistike izvedene iz njih, koriste se za poređenje istih pokazatelja i statistika drugih procesa.

3 Analiza linearnih vremenskih serija

3.1 Linearni procesi

Definicija 13 *Vremenska serija X_t je linearni proces ako se može napisati na sledeći način*

$$X_t = \mu + \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}, \tag{3.1}$$

gde je μ očekivanje slučajne veličine X_t, $\psi_0 = 1$ i ε_t je niz nezavisnih i jednako raspodeljenih slučajnih veličina sa očekivanjem nula i konačnom disperzijom.

Dinamička struktura linearnog procesa opisana je pomoću koeficijenata ψ_i, koji se nazivaju ψ - ponderi.

Kod stacionarnih procesa jedna od bitnih stavki je Voldova [1](Wold) teorema razlaganja, koja govori da se svaki stacionarni proces može izraziti kao zbir dva međusobno nekorelisana procesa, od kojih je jedan deterministički i drugi nedeterministički. Nedeterministički proces se najčešće izražava u formi (3.1), i prema Voldu ova reprezentacija je dovoljno opšta da obuhvati sve slabo stacionarne stohastičke procese.

OSNOVNA SVOJSTVA LINEARNIH PROCESA

Ako je X_t slabo stacionaran slučajan proces, onda važe sledeća svojstva:

1. Matematičko očekivanje

$$E(X_t) = \mu.$$

2. Disperzija

$$\begin{aligned} D(X_t) &= E(X_t - \mu)^2 \\ &= E(\varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \ldots)^2 \\ &= \sigma^2 (1 + \psi_1^2 + \psi_2^2 + \ldots) \\ &= \sigma^2 \sum_{i=0}^{\infty} \psi_i^2. \end{aligned}$$

3. Autokovarijaciona funkcija

$$\begin{aligned} \gamma_k &= cov(X_t, X_{t-k}) \\ &= E(X_t - \mu)(X_{t-k} - \mu) \\ &= \sigma^2 (\psi_k + \psi_1 \psi_{k+1} + \psi_2 \psi_{k+2} + \ldots) \\ &= \sigma^2 \sum_{i=0}^{\infty} \psi_i \psi_{k+i}. \end{aligned}$$

[1] Herman Ole Andreas Wold (1908-1992), švedski matematičar.

4. Autokorelaciona funkcija

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\sum_{i=0}^{\infty}\psi_i\psi_{k+i}}{\sum_{i=0}^{\infty}\psi_i^2}.$$

Stacionarnost linearnih procesa se izražava preko ψ pondera. Njih ima beskonačno mnogo, pa je za stacionarnost procesa neophodno da disperzija bude konačna. Bez umanjenja opštosti, pretpostavimo da je $\mu = 0$, pa imamo da važi

$$|\gamma_k| = |E(X_t X_{t-k})| \le [D(X_t)D(X_{t-k})]^{\frac{1}{2}} = \sigma^2 \sum_{i=0}^{\infty}\psi_j^2.$$

Iz prethodnog sledi da je uslov za stacionarnost linearnog procesa $\sum_{i=0}^{\infty}\psi_j^2 < \infty$.

Na osnovu linearnog procesa (3.1) mogu se izvesti tri klase procesa. Pogodnim izborom pondera ψ dobijamo modele sa konačnim brojem parametara i njima opisujemo slabo stacionarne vremenske serije. Ti modeli su:

1. Autoregresioni model ili AR model,
2. Model pokretnih proseka ili MA model,
3. Autoregresioni model pokretnih proseka ili $ARMA$ model.

3.2 Autoregresioni modeli

Autoregresioni model reda p ili $AR(p)$ definiše se na sledeći način

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + ... + \phi_p X_{t-p} + \varepsilon_t, \tag{3.2}$$

gde su $\phi_1, \phi_2, ..., \phi_p$ autoregresioni parametri, ε_t je proces belog šuma. Ovo je regresioni model, gde zavisna promenljiva predstavljena članom vremenske serije u trenutku t, a objašnjavajuće promenljive su članovi iste vremenske serije samo u trenutcima $t-1,...,t-p$ ([6]).

Autoregresionom modelu reda p pridružuje se karakteristična jednačina koja je data u sledećem obliku:

$$g^p - \phi_1 g^{p-1} - \phi_2 g^{p-2} - ... - \phi_p = 0,$$

gde su $g_1,..., g_p$ koreni - rešenja karakteristične jednačine. Stacionarnost vremenske serije zavisi od rešenja karakteristične jednačine:

• Ako su svi koreni $g_1,...,g_p$ po modulu *strogo manji od jedan*, vremenska serija je stacionarna.

• Ako postoji *bar jedan* koren g_i i $i = 1,2,...,p$ koji je jednak vrednosti jedan po modulu, dok su drugi koreni strogo manji od jedan po modulu, onda je vremenska serija nestacionarna.

• Ako postoji bar jedan koren $g_1,...,g_p$ koji je strogo veći od jedan, dok su drugi strogo manji od jedan po modulu, tada je vremenska serija *eksplozivna*, tj. vremenska serija je pod uticajem aditivnog dejstva trajno rastućeg efekta neočekivanih slučajnih šokova.

3.2.1 Autoregresioni model prvog reda

Autoregresioni model prvog reda ili $AR(1)$ dat je na sledeći način

$$X_t = \phi_1 X_{t-1} + \varepsilon_t \,, \tag{3.3}$$

ϕ_1 je autoregresioni parametar.

• *Uslov stacionarnosti*

Rekurzivnom zamenom unazad dobijamo:

$$\begin{aligned} X_t &= \phi_1 X_{t-1} + \varepsilon_t \\ &= \phi_1[\phi_1 X_{t-2} + \varepsilon_{t-1}] + \varepsilon_t \\ &= \phi_1^2[\phi_1 X_{t-3} + \varepsilon_{t-2}] + \varepsilon_t + \phi_1\varepsilon_{t-1} \\ &= ... = \\ &= \varepsilon_t + \phi_1\varepsilon_{t-1} + \phi_1^2\varepsilon_{t-2} + \phi_1^3\varepsilon_{t-3} + ... \end{aligned}$$

Iz prethodnog računamo disperziju vremenske serije X_t

$D(X_t) = D(\varepsilon_t + \phi_1\varepsilon_{t-1} + \phi_1^2\varepsilon_{t-2} + \phi_1^3\varepsilon_{t-3} + ...) = \sigma^2(1 + \phi_1^2 + \phi_1^4 + \phi_1^6 + ...)$, gde je $\sigma^2 = D(\varepsilon_t)$.

Disperzija vremenske serije je konačna, a sama vremenska serija je slabo stacionarna jedino ako je $|\phi_1| < 1$. Uz ovaj uslov disperzija je jednaka

$$D(X_t) = \frac{\sigma^2}{1 - \phi_1^2}.$$

• *AR(1) model je specijalan slučaj linearnog procesa*

Korišćenjem operatora L, gde je L operator pomeraja, $L(X_t) = X_{t-1}$, model se može zapisati na sledeći način:

$$(1 - \phi_1 L)X_t = \varepsilon_t \Rightarrow X_t = \frac{1}{1 - \phi_1 L}\varepsilon_t.$$

Za $|\phi_1|<1$, $\dfrac{1}{(1-\phi_1 L)}$ predstavlja zbir članova geometrijskog reda $(1+\phi_1 L+\phi_1^2 L^2+\phi_1^3 L^3+...)$, tako da važi:

$$\begin{aligned} X_t &= \frac{1}{1-\phi_1 L}\varepsilon_t \\ &= (1+\phi_1 L+\phi_1^2 L^2+\phi_1^3 L^3+...)\varepsilon_t \\ &= \varepsilon_t+\psi_1\varepsilon_{t-1}+\psi_2\varepsilon_{t-2}+\psi_3\varepsilon_{t-3}+... \end{aligned}$$

gde je $\psi_1=\phi_1, \psi_2=\phi_1^2,...$

Iz prethodnog vidimo da $AR(1)$ za $|\phi_1|<1$ jeste specijalan slučaj linearnog procesa.

- *Autokovarijaciona i autokorelaciona funkcija*

Za $E(X_t)=0$, autokovarijacioni koeficijent na rastojanju k je $\gamma_k=E(X_t X_{t-k})$.

Kako je

$$X_t=\varepsilon_t+\phi_1\varepsilon_{t-1}+\phi_1^2\varepsilon_{t-2}+\phi_1^3\varepsilon_{t-3}+...+\phi_1^k\varepsilon_{t-k}+\phi_1^{k+1}\varepsilon_{t-k-1}+...,$$

to važi da je i

$$X_{t-k}=\varepsilon_{t-k}+\phi_1\varepsilon_{t-k-1}+\phi_1^2\varepsilon_{t-k-2}+...,$$

pa je autokovarijacioni koeficijent na rastojanju k

$$\begin{aligned} \gamma_k &= E(X_t X_{t-k}) \\ &= E((\varepsilon_t+\phi_1\varepsilon_{t-1}+\phi_1^2\varepsilon_{t-2}+...+\phi_1^{k+1}\varepsilon_{t-k-1}+...)(\varepsilon_{t-k}+\phi_1\varepsilon_{t-k-1}+\phi_1^2\varepsilon_{t-k-2}+...)) \\ &= \sigma^2(\phi_1^k+\phi_1\phi_1^{k+1}+\phi_1^2\phi_1^{k+2}+...) \\ &= \sigma^2\phi_1^k(1+\phi_1^2+\phi_1^4+...) \\ &= \frac{\sigma^2\phi_1^k}{1-\phi_1^2} \end{aligned}$$

pri rastojanju k autokovarijaciona funkcija zavisi od disperzije belog šuma, i autoregresionog parametra.

Autokorelacioni koeficijent na rastojanju k dat je sledećom formulom $\rho_k=\dfrac{\gamma_k}{\gamma_0}$.

Koristeći prethodne formule dobijamo:

$$\rho_k=\frac{\dfrac{\sigma^2\phi_1^k}{1-\phi_1^2}}{\dfrac{\sigma^2}{1-\phi_1^2}}=\phi_1^k.$$

Pa je autokorelaciona funkcija vremenske serije opisane AR(1) modelom jednaka:

$$\rho_k=\phi_1^k, k=1,2,...$$

3.2.2 Autoregresioni model reda p

Za autoregresioni modela reda p ili $AR(p)$ koji je dat u sledećem obliku

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + ... + \phi_p X_{t-p} + \varepsilon_t ,$$

gde su $\phi_1, \phi_2, ..., \phi_p$ autoregresioni parametri, dok je ε_t beli šum, važi:

- *Uslov stacionarnosti*

$AR(p)$ modelu odgovara diferencna jednačina reda p

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - ... - \phi_p X_{t-p} = \varepsilon_t ,$$

čija je karakteristična jednačina

$$g_p - \phi_1 g_{p-1} - \phi_2 g_{p-2} - ... - \phi_p = 0.$$

Ukoliko su rešenja ove jednačine po modulu manja od jedan, vremenska serija je stacionarna.

- *Autokovarijaciona i autokorelaciona funkcija*

Autokovarijaciona funkcija autoregresionog modela red p dobija se na sledeći način

$$\gamma_k = \begin{cases} \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + ... + \phi_p \gamma_{k-p} + \sigma^2 & , za \quad k = 0 \\ \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + ... + \phi_p \gamma_{k-p} & , za \quad k > 0 \end{cases}$$

Za $k = 0$ autokovarijaciona funkcija se svodi na disperziju vremenske serije

$$\gamma_0 = \frac{\sigma^2}{1 - \phi_1 \rho_1 - \phi_2 \rho_2 - ... - \phi_p \rho_p}.$$

Kada je $k > 0$ dobijamo

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + ... + \phi_p \gamma_{k-p}.$$

Deljenjem sa γ_0 dobijamo autokorelacionu funkciju autoregresionog modela reda p

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + ... + \phi_p \rho_{k-p}, za \quad k > 0.$$

3.3 Modeli pokretnih proseka

Model pokretnih proseka reda q ili $MA(q)$ definiše se na sledeći način:

$$X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - ... - \theta_q \varepsilon_{t-q} , \qquad (3.4)$$

gde je ε_t proces belog šuma, dok su $\theta_1, \theta_2, ..., \theta_q$ parametri modela.

Nivo vremenske serije u trenutku t opisuje se u funkciji od članova procesa belog šuma u trenucima $t, t-1, ..., t-q$. Pomeranjem u vremenu menjaju se neki članovi niza ali broj sabiraka ostaje isti. Može se govoriti o izravnanju prilikom definisanja vremenske serije, jer ovako obrazovani zbirovi sadrže veći broj sabiraka sa sličnim uticajem ([6]).

Modelom $MA(q)$ definisana je slabo stacionarna vremenska serija. To vidimo iz činjenice da se model sastoji od konačnog zbira članova belog šuma na različitim rastojanjima. Iz prethodnog, takođe, sledi i da je disperzija vremenske serije konačna

$$D(X_t) = D(\varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - ... - \theta_q\varepsilon_{t-q}) = \sigma^2(1+\theta_1^2+\theta_2^2+...+\theta_q^2).$$

Ako uvedemo smenu $\psi_1 = -\theta_1,\ \psi_2 = -\theta_2, ..., \psi_q = -\theta_q,\ \psi_j = 0, j > q$ model $MA(q)$ možemo svesti na formu linearnog procesa

$$X_t = \varepsilon_t + \psi_1\varepsilon_{t-1} + \psi_2\varepsilon_{t-2} + ...$$

Linearni proces može se označiti i kao MA model beskonačnog reda, $MA(\infty)$.

3.3.1 Model pokretnih proseka prvog reda

Model pokretnih proseka prvog reda ili $MA(1)$ dat je na sledeći način

$$X_t = \varepsilon_t - \theta_1\varepsilon_{t-1}, \tag{3.5}$$

θ_1 je parametar modela.

- *Autokovarijaciona funkcija*

Kako je

$$\gamma_k = E(X_t X_{t-k}) = E(\varepsilon_t - \theta_1\varepsilon_{t-1})(\varepsilon_{t-k} - \theta_1\varepsilon_{t-k-1}),$$

dobija se autokovarijaciona funkcija

$$\gamma_k = \begin{cases} (1+\theta_1^2)\sigma^2 & , k = 0 \\ -\theta_1\sigma^2 & , k = 1 \\ 0 & , k > 1 \end{cases}$$

- *Autokorelaciona funkcija*

Autokorelacioni koeficijent na rastojanju k dat je sledećom formulom $\frac{\gamma_k}{\gamma_0}$. Na osnovu prethodne formule za autokovarijacionu funkciju, dobijamo da je autokorelaciona funkcija data sledećim izrazom:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \begin{cases} 1 & ,k=0 \\ -\dfrac{\theta_1}{1+\theta_1^2} & ,k=1 \\ 0 & ,k>1 \end{cases}$$

Autokorelaciona funkcija $MA(1)$ modela ima dva bitna svojstva:

(a) Ne postoji jednoznačna određenost obične autokorelacione funkcije. Na primer, istu korelacionu funkciju poseduju sledeća dva modela

$$X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1}, \quad \rho = -\frac{\theta_1}{1+\theta_1^2} \text{ i}$$

$$X_t = \varepsilon_t - \frac{1}{\theta_1}\varepsilon_{t-1}, \quad \rho = -\frac{\frac{1}{\theta_1}}{1+\frac{1}{\theta_1^2}} = -\frac{\theta_1}{1+\theta_1^2}.$$

(b) Autokorelacioni koeficijent uzima vrednosti iz intervala (-0.5,0.5), što je uži interval od onog za koji je definisana autokorelaciona funkcija, tj. (-1,1).

- *Uslov invertibilnosti*

Možemo videti da važi

$$X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} \Rightarrow \varepsilon_t = X_t + \theta_1 \varepsilon_{t-1} \Rightarrow \begin{cases} \varepsilon_{t-1} = X_{t-1} + \theta_1 \varepsilon_{t-2} \\ \varepsilon_{t-2} = X_{t-2} + \theta_1 \varepsilon_{t-3} \\ itd. \end{cases}$$

Kada ubacimo ovo u polazni model dobijamo

$$\begin{aligned} X_t &= \varepsilon_t - \theta_1 \varepsilon_{t-1} \\ &= \varepsilon_t - \theta_1 (X_{t-1} + \theta_1 \varepsilon_{t-2}) \\ &= -\theta_1 X_{t-1} - \theta_1^2 (X_{t-2} + \theta_1 \varepsilon_{t-3}) + \varepsilon_t \\ &= \ldots = \\ &= -\theta_1 X_{t-1} - \theta_1^2 X_{t-2} - \theta_1^3 X_{t-3} - \ldots + \varepsilon_t. \end{aligned}$$

Uvodimo smenu: $\pi_j = -\theta_1^j,\ j = 1,2,\ldots$ i dobijamo autoregresioni model beskonačnog reda

$$X_t = \pi_1 X_{t-1} + \pi_2 X_{t-2} + \pi_3 X_{t-3} + \ldots + \varepsilon_t.$$

Potrebno je odrediti pod kojim uslovima dati model opisuje stacionarnu vremensku seriju. Na osnovu toga kako su definisani parametri modela, a da bi smo dobili da je serija stacionarna zaključujemo da parametar θ_1 treba da je po modulu manji od jedan. Time smo dobili uslov invertibilnosti, gde se podrazumeva ekvivalentnost $MA(1)$ modela i odgovarajućeg autoregresionog modela beskonačnog reda.

Vremenske serije koje zadovoljavaju uslov invertibilnosti su invertibilne vremenske serije.

3.3.2 Model pokretnih proseka reda q

Za model pokretnih proseka reda q ili $MA(q)$ koji ima sledeći oblik

$$X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - ... - \theta_q \varepsilon_{t-q}$$

važi

- *Autokovarijaciona funkcija*

$$\gamma_k = \begin{cases} (1+\theta_1^2 + ... + \theta_q^2)\sigma^2 & ,k = 0 \\ (-\theta_k + \theta_1\theta_{k+1} + ... + \theta_{q-k}\theta_q)\sigma^2 & ,k = 1,2,...,q \\ 0 & ,k > q \end{cases}$$

- *Autokorelaciona funkcija*

$$\rho_k = \begin{cases} 1 & ,k = 0 \\ \dfrac{(-\theta_k + \theta_1\theta_{k+1} + ... + \theta_{q-k}\theta_q)}{1+\theta_1^2 + ... + \theta_q^2} & ,k = 1,2,...,q \\ 0 & ,k > q \end{cases}$$

3.4 Autoregresioni model pokretnih proseka

Često se u radu sa MA i AR modelima javlja potreba za velikim brojem parametara. Međutim, veliki broj parametara otežava rad. Da bi se rešio ovaj problem, uveden je novi model za opisivanje vremenskih serija. Taj model se naziva autoregresivni model pokretnih proseka ili $ARMA$, i on je kombinacija AR i MA modela, kao što i samo ime kaže ([6]).

3.4.1 ARMA(p,q)

Autoregresioni modeli pokretnih proseka ili $ARMA$ dati su sledećom formulom

$$\underbrace{(1-\phi_1 L - \phi_2 L^2 - ... - \phi_p L^p)}_{\Phi(L)} X_t = \underbrace{(1-\theta_1 L - \theta_2 L^2 - ... - \theta_q L^q)}_{\Theta(L)} e_t. \tag{3.6}$$

Oznaka koju koristimo za dati model je $ARMA(p,q)$, gde je

- p - red autoregresione komponente
- q - red komponente pokretnih proseka

L je operator pomeraja $L(X_t) = X_{t-1}$, a za više stepene rekurzivnom vezom dobijamo $L^k(X_t) = L^{k-1}(X_{t-1})$.

Za polinome $\Phi(L)$ i $\Theta(L)$ se pretpostavlja da ne sadrže zajedničke faktore i da opisuju redom autoregresionu i komponentu pokretnih proseka stacionarne i invertibilne vremenske serije X_t. Proces e_t je beli šum.

Specijalni slučajevi ARMA(p,q) modela:

AR(p) = ARMA(p,0)
MA(q) = ARMA(0,q)

- *Stacionarnost*

Stacionarnost ovako definisane vremenske serije određena je stacionarnošću odgovarajuće autoregresione komponente.

- *Autokovarijaciona funkcija*

Autokovarijaciona funkcija jednaka

$$\begin{aligned}\gamma_k &= E(X_t X_{t-k})\\ &= \phi_1\gamma_{k-1} + \phi_2\gamma_{k-2} + \dots + \phi_p\gamma_{k-p} + E(\varepsilon_t X_{t-k}) - \theta_1 E(\varepsilon_{t-1}X_{t-k})\\ &\quad - \theta_2 E(\varepsilon_{t-2}X_{t-k}) - \dots - \theta_q E(\varepsilon_{t-q}X_{t-k}).\end{aligned}$$

Kako je data vremenska serija linearan proces, zaključujemo da važi

$$E(\varepsilon_{t-q}X_{t-k}) = \begin{cases} 0 & ,k > q \\ D(\varepsilon_t) = \sigma^2 & ,k = q \end{cases}$$

Prema tome, za $k > q$ autokovarijaciona funkcija je

$$\gamma_k = \phi_1\gamma_{k-1} + \phi_2\gamma_{k-2} + \dots + \phi_p\gamma_{k-p}.$$

- *Autokorelaciona funkcija*

Autokorelaciona funkcija ima sledeći oblik

$$\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2} + \dots + \phi_p\rho_{k-p}, za \quad k > q.$$

Donosimo sledeće zaključke vezane za svojstva autokorelacione funkcije:

(a) Prvih q autokorelacionih koeficijenata vremenske serije generisane $ARMA(p,q)$ modelom zavise od parametara AR i MA komponente.

(b) Za rastojanja koja su veća od reda MA komponente, $q+1$, $q+2$, itd. ova funkcija ponaša se kao autokorelaciona funkcija AR modela.

3.5 Autoregresioni model pokretnih proseka za integrisane serije

Definicija 14 *Operacija oduzimanja uzastopnih opservacija* $X_t - X_{t-1} = \Delta X_t$ *naziva se diferenciranje, a operator* Δ *je diferencni operator.*

Definicija 15 *Stohastički proces* $\{X_t, t \in T\}$ *je integrisan proces reda d u oznaci* I (d) *ako može biti transformisan u stacionaran stohastički proces diferenciranjem d puta.*

Često se u radu sa ARMA modelima javlja problem stacionarnosti. Međutim problem se pravazilazi uvođenjem ARIMA modela. ARIMA model nestacionarnu vremensku seriju svodi prvo na stacionarnu, diferenciranjem, a zatim seriju modelira ARMA modelom.

3.5.1 ARIMA(p,d,q)

Autoregresioni model pokretnih proseka za integrisane vremenske serije opisuju se formulom oblika

$$\underbrace{(1-\phi_1 L - \phi_2 L^2 - ... - \phi_p L^p)}_{\Phi(L)}(1-L)^d X_t = \underbrace{(1-\theta_1 L - \theta_2 L^2 - ... - \theta_q L^q)}_{\Theta(L)} e_t. \quad (3.7)$$

Za proces (u našem slučaju vremensku seriju) X_t kaže se da je reda integrisanosti d, u oznaci I(d). U realnim primenama, kod ekonomskih vremenskih serija, obično je broj potrebnih razlika - diferenciranja d jednak 0, 1 ili 2. Procesi kod kojih se stacionarnost postiže diferenciranjem nazivaju se diferencno stacionarni procesi. Nekada je potrebno više puta diferencirati proces da bi se postigla stacionarnost ([6]).

Oznaka koju koristimo za dati model je $ARIMA(p,d,q)$, gde je

- p - red autoregresione komponente
- d - nivo integrisanosti vremenske serije
- q - red autoregresione komponente

L je operator pomeraja $L(X_t) = X_{t-1}$, a za više stepene rekurzivnom vezom dobijamo $L^k(X_t) = L^{k-1}(X_{t-1})$.

Za polinome $\Phi(L)$ i $\Theta(L)$ se pretpostavlja da ne sadrže zajedničke faktore i da opisuju redom autoregresionu i komponentu pokretnih proseka stacionarne i invertibilne vremenske serije $(1-L)^d X_t$. Proces e_t je beli šum.

Specijalni slučajevi ARIMA(p,d,q) modela:

AR(p) = ARMA(p,0,0)
MA(q) = ARMA(0,0,q)
beli šum = ARIMA(0,0,0)
slučajni hod = ARIMA(0,1,0)

Pretpostavimo da je d=1. U tom slučaju imao ARIMA(p,1,q) model oblika:

$$\Delta X_t = \theta_0 + \phi_1 \Delta X_{t-1} + \phi_2 \Delta X_{t-2} + ... + \phi_p \Delta X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - ... - \theta_q \varepsilon_{t-q} .$$

Reč je o ARIMA(p,d,q) modelu, ali za vremensku seriju ΔX_t koja je prethodno transformisana da bi se postigla njena stacionarna reprezentacija. Parametar θ_0 označava konstantni prirast (ili pad), odnosno parametar uz linearni trendu nivou date vremenske serije.

Za d=2 relevantan je model ARIMA(p,2,q) oblika:

$$\Delta^2 X_t = \theta_0 + \phi_1 \Delta^2 X_{t-1} + \phi_2 \Delta^2 X_{t-2} + ... + \phi_p \Delta^2 X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - ... - \theta_q \varepsilon_{t-q}.$$

Vremenska serija koja se ovde opisuje posedujedva jedinična korena. Kako je njena druga diferenca stacionarna, to se ARMA formom modelira upravo vremenska serija formirana kao druga diferenca polazne serije. U ovom modelu parameter θ_0 može da se dovede u vezu sa parametrom uz kvadratni trend u nivou serije ΔX_t.

4 Prognoziranje budućih vrednosti - Modeli

Jako važan zadatak analize vremenskih serija je predviđanje - prognoziranje, tj. određivanje budućeg toka posmatrane vremenske serije. Bitne odluke u političkom, poslovnom i svakodnevnom životu se donose na osnovu prognoziranih podataka, pa je jako bitno da prognoza bude što tačnija, jer u suprotnom pogrešna odluka može dovesti do velikih materijalnih gubitaka. U procesu prognoziranja imamo vremensku seriju za koju su poznati podaci do trenutka h, $\mathbf{X} = (X_1, ..., X_h)$, i potrebno je da odredimo buduće vrednosti za ℓ koraka unapred, tj. $X_{h+\ell}$, gde je $\ell \geq 1$. Broj ℓ se naziva horizont predviđanja. Kod mesečnih serija pogodno je prognozirati za tri koraka unapred (buduća tri meseca), kod kvartalnih jedan (eventualno dva) sledeća kvartala, dok kod dnevnih se prognozira i na duži vremenski period (više perioda - dana pa čak i meseci), ali je diskutabilno koliko su te prognoze merodavne npr. vremenska prognoza.

Sve preko pomenutih ograničenja, već ne bi bile prognoze, nego projekcije (prognoza za više od tri meseca unapred – kod mesečnih serija, za više od kvartala – kod kvartalih itd.). Projekcije iziskuju dodatna razmatranja i uključiuju u “igru” još dodatnih faktora i istraživanja. Može se reći da je svaka prognoza ujedno i projekcija, ali da svaka projekcija nije prognoza. To znači da je termin projekcija širi po značenju i obuhvata prognozu. Projekcije daju objašnjenje o tome kakva bi serija bila ako bi se izvesne pretpostavke ostvarile. One se, međutim, ne bave i pitanjem da li će se one i ostvariti. Projekcije su zbog toga uvek "ispravne", dok prognoza može biti i pogrešna.

U ovoj glavi bavićemo se sledećim modelima za prognoziranje: *prognouziranje sa minimalnom srednje kvadratnom greškom*, *prognoziranje budućih vrednosti preko AR(p) modela*, *prognoziranje budućih vrednosti preko MA(q) modela*, *prognoziranje budućih vrednosti preko ARMA(p,q) modela*, *prognoziranje budućih vrednosti koristeći eksponencijalno izravnanje*, *prognoziranje budućih vrednosti preko ARIMA(p,d,q) modela*, *prognoziranje budućih vrednosti preko sezonskih ARIMA modela*.

4.1 Prognoziranje sa minimalnom srednje kvadratnom greškom

Kod formiranja prognoze, cilj nam je da dobijeni rezultati imaju što manju grešku, tj. da budu što bliži stvarnoj vrednosti, vrednosti koja će se ostvariti. Iz tog razloga se formira prognoza koja minimizuje srednje kvadratnu grešku. Ako je vremenska serija stacionarna, onda promenljive $X_1,...,X_h$ imaju konstantno očekivanje μ. Kao prirodno rešenje, nameće se korišćenje sredine, kao prognoze budućih vrednosti. Osim toga što je logična, ova prognoza je i optimalna, jer se može pokazati da među svim prognozama ona ima najmanju srednje kvadratnu grešku. Međutim, korišćenje očekivanja kao prognoze budućih vrednosti i nije baš najadekvatnije, jer početak i kraj intervala (prvi i poslednji meseci u godini - ako se radi o mesečnim podacima) ili preceni ili podceni.

Bolju prognozu dobijamo korišćenjem raspoloživih informacija, pa tako ćemo umesto funkcije gustine promenljive $X_{h+\ell}$, koristiti uslovnu funkciju gustine, uslovljenu poznatim vrednostima. Takođe, i srednje kvadratnu grešku ćemo formirati kao uslovnu očekivanu vrednost, uslovljenu poznatim informacijama.

S ciljem da izvedemo prognozu minimalne srednje kvadratne greške, polazimo od ARMA modela. Izrazićemo ARMA model u MA formi preko izraza za linearni proces

$$X_t = \varepsilon_t + \psi_1\varepsilon_{t-1} + \psi_2\varepsilon_{t-2} + ...$$

zamenom $t = h + \ell$, dobijamo

$$X_{h+\ell} = \sum_{j=0}^{\infty}\psi_j\varepsilon_{h+\ell-j}.$$

Poznate su nam vrednosti do trenutka h, i potrebno je da odredimo buduće vrednosti za ℓ koraka unapred i to u obliku linearne kombinacije promenljivih $X_h, X_{h-1}, X_{h-2},...$.

Prognoza $\hat{X}_h(\ell)$ za $X_{h+\ell}$ je data izrazom

$$\hat{X}_h(\ell) = \psi^*_\ell\varepsilon_h + \psi^*_{\ell+1}\varepsilon_{h-1} + \psi^*_{\ell+2}\varepsilon_{h-2} + ...$$

gde treba odrediti koeficijente ψ^*_j. Vrednost ovih koeficijenata dobijamo minimizovanjem srednje kvadratne greške. Greška prognoze je

$$X_{h+\ell} - \hat{X}_h(\ell) = \left(\sum_{j=0}^{\ell-1}\psi_j\varepsilon_{h+\ell-j} + \sum_{j=0}^{\infty}\psi_{\ell+j}\varepsilon_{h-j}\right) - \sum_{j=0}^{\infty}\psi^*_{\ell+j}\varepsilon_{h-j}$$

pa je srednje kvadratna greška prognoze

$$E(X_{h+\ell} - \hat{X}_h(\ell))^2 = \sigma^2\sum_{j=0}^{\ell-1}\psi_j^2 + \sigma^2\sum_{j=0}^{\infty}(\psi_{\ell+j} - \psi^*_{\ell+j})^2.$$

Očigledno je da se minimalana vrednost srednje kvadratne greške postiže kada je $\psi^*_{\ell+j} = \psi_{\ell+j}$, pa je prognoza sa minimalnom srednje kvadratnom greškom

$$\hat{X}_h(\ell) = \psi_\ell \varepsilon_h + \psi_{\ell+1}\varepsilon_{h-1} + \psi_{\ell+2}\varepsilon_{h-2} + ...$$

$\hat{X}_h(\ell)$ za $\ell = 1,2,...$ naziva se funkcija prognoze. Uslovna očekivana vrednost šokova je

$$E(\varepsilon_{h+j} \mid X_h, X_{h=1},...) = \begin{cases} \varepsilon_{h+j}, & j \le 0 \\ 0, & j > 0 \end{cases}$$

Iz prethodnog dalje dobijamo

$$E(X_{h+\ell} \mid X_h, X_{h-1},...) = \psi_\ell \varepsilon_h + \psi_{\ell+1}\varepsilon_{h-1} + \psi_{\ell+2}\varepsilon_{h-2} + ...$$

što zapravo znači da je prognoza za $X_{h+\ell}$ data njenom uslovnom očekivanom vrednosti

$$\hat{X}_h(\ell) = E(X_{h+\ell} \mid X_h, X_{h-1},...).$$

Može se pokazati da je prognoza nepristrasna ocena buduće vrednosti. To sledi na osnovu greške prognoze za ℓ perioda unapred

$$e_h(\ell) = X_{h+\ell} - \hat{X}_h(\ell) = \sum_{j=0}^{\ell-1} \psi_j \varepsilon_{h+\ell-j}$$

prema kojoj je $E(e_h(\ell) \mid X_t, t \le h) = 0$. Disperzija greške prognoze za ℓ perioda unapred je

$$D(\hat{X}_h(\ell)) = D(e_h(\ell)) = \sigma^2 \sum_{j=0}^{\ell-1} \psi_j^2.$$

4.2 Prognoziranje budućih vrednosti preko AR(p) modela

U $AR(p)$ modelu koji je dat jednačinom (3.2) pretpostavimo da su nam poznate vrednosti do trenutka h, naš zadatak je da prognoziramo vrednosti za ℓ koraka unapred. Sa $\hat{X}_h(\ell)$ označimo prognozu za $X_{h+\ell}$ dobijenu korišćenjem funkcije minimalne srednje kvadratne greške i F_h kolekcija poznatih informacija do trenutka h. $\hat{X}_h(\ell)$ biramo tako da

$$E\{[X_{h+\ell} - \hat{X}_h(\ell)]^2 \mid \mathsf{F}_h\} \le \min_g E[(X_{h+\ell} - g)^2 \mid \mathsf{F}_h]$$

gde je g funkcija poznatih promenljivih do vremena h.

PROGNOZA ZA JEDAN KORAK UNAPRED

Iz AR(p) modela imamo

$$X_{h+1} = \phi_0 + \phi_1 X_h + \phi_2 X_{h-1} + ... + \phi_p X_{h-p+1} + \varepsilon_{h+1}$$

gde je $\phi_0 = (1-\phi_1)\mu, E(X_t) = \mu$.

Na osnovu principa minimalne srednje kvadratne greške, prognoziranje vrednosti X_{h+1} se vrši preko uslovnog očekivanja

$$\hat{X}_h(1) = E(X_{h+1} \mid F_h) = \phi_0 + \sum_{i=1}^{p} \phi_i X_{h+1-i}.$$

Greška prognoze je

$$e_h(1) = X_{h+1} - \hat{X}_h(1) = \varepsilon_{h+1}.$$

Stoga, disperzija greške prognoze za jedan korak je $D(e_h(1)) = D(\varepsilon_{h+1}) = \sigma_\varepsilon^2$. Ako ε_t ima normalnu raspodelu, tada je interval poverenja za X_{t+1} jednak $\hat{X}_h(1) \pm 1.96 \times \sigma_\varepsilon$, sa novoom poverenja 95%.

PROGNOZA ZA DVA KORAKA UNAPRED

Sada računamo za dva koraka unapred. Iz $AR(p)$ modela imamo

$$X_{h+2} = \phi_0 + \phi_1 X_{h+1} + \phi_2 X_h + \ldots + \phi_p X_{h-p+2} + \varepsilon_{h+2}.$$

Uzimajući uslovno očekivanje, imamo

$$\hat{X}_h(2) = E(X_{h+2} \mid F_h) = \phi_0 + \phi_1 \hat{X}_h + \phi_2 X_h + \ldots + \phi_p X_{h+2-p}.$$

Greška prognoziranja za dva koraka

$$e_h(2) = X_{h+2} - \hat{X}_h(2) = \phi_1[X_{h+1} - \hat{X}_h(1)] + \varepsilon_{h+2} = \varepsilon_{h+2} + \phi_1 \varepsilon_{h+1}.$$

Disperzija greške prognoziranja za dva koraka je $D(e_h(2)) = (1+\phi_1^2)\sigma_\varepsilon^2$. Interval poverenja se dobija na sličan način kao za jedan korak unapred.

PROGNOZA ZA ℓ KORAKA UNAPRED

Polazimo od

$$X_{h+\ell} = \phi_0 + \phi_1 X_{h+\ell-1} + \ldots + \phi_p X_{h-p+\ell} + \varepsilon_{h+\ell}.$$

Uzimajući uslovno očekivanje $\hat{X}_h(l) = E(X_{h+l} \mid F_h)$, imamo

$$\hat{X}_h(\ell) = \phi_0 + \sum_{i=1}^{p} \phi_i \hat{X}_h(\ell - i).$$

Greška prognoziranja za ℓ koraka

$$e_h(\ell) = X_{h+\ell} - \hat{X}_h(\ell).$$

Može se pokazati da za stacionaran $AR(p)$ model, $\hat{X}_h(\ell)$ konvergira ka $E(X_t)$ kad $\ell \to \infty$. Iz ovog vidimo da se dugoročna prognoza približava bezuslovnom očekivanju.

4.3 Prognoziranje budućih vrednosti preko MA(q) modela

Kao što smo i kod $AR(p)$, pretpostavimo da su nam poznate vrednosti do trenutka h, naš zadatak je da prognoziramo vrednosti za ℓ koraka unapred. $\hat{X}_h(\ell)$ je prognozirana vrednost za $X_{h+\ell}$ dobijenu korišćenjem funkcije minimalne srednje kvadratne greške i F_h kolekcija poznatih informacija do trenutka h.

PROGNOZA ZA JEDAN KORAK UNAPRED

Posmatramo $MA(1)$ model

$$X_{h+1} = c_0 + \varepsilon_{h+1} - \theta_1 \varepsilon_h,$$

gde je c_0 očekivanje od X_t koje je jednako $c_0 = E(X_t)$.

Korišćenjem uslovnog očekivanja imamo

$$\hat{X}_h(1) = E(X_{h+1} \mid F_h) = c_0 - \theta_1 \varepsilon_h.$$

Greška prognoze je

$$e_h(1) = X_{h+1} - \hat{X}_h(1) = \varepsilon_{h+1}.$$

Disperzija greške prognoze za jedan korak je $D(e_h(1)) = D(\varepsilon_{h+1}) = \sigma_\varepsilon^2$.

PROGNOZA ZA DVA KORAKA UNAPRED

Sada računamo dalje za dva koraka unapred. Iz $MA(1)$ modela imamo

$$X_{h+2} = c_0 + \varepsilon_{h+2} - \theta_1 \varepsilon_{h+1}$$

tj.

$$\hat{X}_h(2) = E(X_{h+2} \mid F_h) = c_0.$$

Greška prognoziranja za dva koraka

$$e_h(2) = X_{h+2} - \hat{X}_h(2) = \varepsilon_{h+2} - \theta_1 \varepsilon_{h+1}.$$

Disperzija greške prognoziranja za dva koraka je $D(e_h(2)) = (1+\theta_1^2)\sigma_\varepsilon^2$. Disperzija za dva koraka je veća ili jednaka od disperzije za jedan korak. Rezultati pokazuju da prognoza za dva koraka unapred je bezuslovno očekivanje modela. Generalno, imamo da važi $\hat{X}_\ell = c_0$ za $\ell \geq 2$. Zaključujemo da je prognozirana vrednost kod $MA(1)$ modela ista za $h > 1$.

Slično za $MA(2)$ modela dobijamo sledeće:

$$\hat{X}_h(1) = c_0 - \theta_1\varepsilon_h - \theta_2\varepsilon_{h-1}$$

$$\hat{X}_h(2) = c_0 - \theta_2\varepsilon_h$$

$$\hat{X}_h(\ell) = c_0, \quad za \quad \ell > 2.$$

Kod $MA(2)$ modela imamo da se vrednosti ponavljaju nakon dva koraka. Uopšteno kod $MA(q)$ modela, prognozirane vrednosti se ponavljaju nakon q koraka.

4.4 Prognoziranje budućih vrednosti preko ARMA(p,q) modela

Karakteristike ARMA(p,q) modela, zavise od parametara modela. Ako je $p > q$, ARIMA(p,q= model se ponaša kao AR(p) i obratno.

PROGNOZA ZA JEDAN KORAK UNAPRED

$$\hat{X}_h(1) = E(X_{h+1} \mid \mathrm{F}_h) = \phi_0 + \sum_{i=1}^{p}\phi_i X_{h+1-i} - \sum_{i=1}^{q}\theta_i\varepsilon_{h+1-i}.$$

Greška prognoze je $e_h(1) = X_{h+1} - \hat{X}_h(1) = \varepsilon_{h+1}$. Disperzija greške prognoze za jedan korak je $D(e_h(1)) = \sigma_\varepsilon^2$.

PROGNOZA ZA ℓ KORAKA UNAPRED

$$\hat{X}_h(\ell) = E(X_{h+\ell} \mid \mathrm{F}_h) = \phi_0 + \sum_{i=1}^{p}\phi_i \hat{X}_h(\ell - i) - \sum_{i=1}^{q}\theta_i\varepsilon_h(\ell - i),$$

gde je $\hat{X}_h(\ell - i) = X_{h+\ell-i}$ ako je $\ell - i \leq 0$, $\varepsilon_h(\ell - i) = 0$ ako je $\ell - i > 0$ i $\varepsilon_h(\ell - i) = \varepsilon_{h+\ell-i}$ ako je $\ell - i \leq 0$.

Greška prognoziranja za ℓ koraka je

$$e_h(\ell) = X_{h+\ell} - \hat{X}_h(\ell).$$

4.5 Prognoziranje budućih vrednosti koristeći eksponencijalno izravnanje

Metoda pokretnih proseka nije najbolja za prognoziranje, jer je usmerena ka središtu posmatranog intervala, dok je za prognoze najbitniji kraj interval. Iz pomenutog razloga uvedeni su ponderisani pokretni proseci čiji ponderi opadaju eksponencijalno, sa uključivanjem ranijih opservacija u model. Naime, smatra se da je od najvećeg uticaja prethodni član zatim onaj pre njega itd. Matematička formulacija može da pođe od sledećeg. Neka je data vremenska serija $X_{1,\dots,}X_h$ i neka se sa $\hat{X}_h(1)$ označe njene izravnate vrednosti, odnosno prognoze . ([10]) Relacija

$$\hat{X}_h(1) = aX_h + (1-a)\hat{X}_h \,, \quad 0 < a < 1 \tag{4.1}$$

uspostavlja vezu između prognoza i članova serije i pogodna je za uključivanje novih opservacija jer je nova prognoza funkcija samo zadnje observacije i njene prognoze. Opravdanost naziva eksponencijalno izravnanje vidi se tek sukcesivnom primenom jednakosti (4.1) čime se dobija

$$\hat{X}_h(1) = aX_h + (1-a)(aX_{h-1} + (1-a)\hat{X}_{h-1}) = aX_h + a(1-a)X_{h-1} + (1-a)^2\hat{X}_{h-1}$$

$$= aX_h + a(1-a)X_{h-1} + (1-a)^2(aX_{h-2} + (1-a)\hat{X}_{h-2})$$

$$= \dots =$$

$$= aX_h + a(1-a)X_{h-1} + a(1-a)^2 X_{h-2} + \dots + a(1-a)^{n-1}X_{h-n+1}$$

Ponderi (koeficijenti uz X_h) eksponencijalno opadaju tako da je svaki naredni $(1-a)$ puta manji od prethodnog. Najveći ponder, jednak a, odgovara najnovijoj opservaciji. Drugi oblik modela eksponencijalnog izravnanja uključuje grešku prognoze e_t u izraz (4.1)

$$e_h = X_h - \hat{X}_h$$

$$\hat{X}_h(1) = aX_h + (1-a)\hat{X}_h$$

$$= \hat{X}_h + a(X_h - \hat{X}_h)$$

$$= \hat{X}_h + ae_h .$$

4.6 Prognoziranje budućih vrednosti preko ARIMA(p,d,q) modela

Definicija 16 $\{X_t, t \in T\}$ *je proces slučajnog hoda ako važi*:

$$X_t = X_{t-1} + \varepsilon_t,$$

gde je $\{\varepsilon_t, t \in T\}$ proces belog šuma.

Slučajni hod je dobio ime zbog karakterističnog ponašanja koje podseća na putanju kretanja pijanog čoveka koji slučajno krivuda. Vremensku seriju koja prati slučajan hod karakterišu periodi rastućeg trenda, gde se trend kretanja menja iznenada i promena je nepredvidiva.

Za proces slučajnog hoda važi: $X_{t-1} = X_{t-2} + \varepsilon_{t-1}, ..., X_1 = X_0 + \varepsilon_1$.
Dakle:

$$X_t = X_o + \sum_{i=1}^{t} \varepsilon_i,$$

odakle sledi da je $D(X_t) = D(X_0 + \sum_{i=1}^{t} \varepsilon_i) = t\sigma^2$. Kako disperzija raste tokom vremena, slučajan hod je nestacionaran proces.
Ukoliko se sa obe strane jednakosti oduzme X_{t-1} dobija se stacionaran proces:

$$X_t = X_{t-1} + \varepsilon_t \,/ - X_{t-1} \Rightarrow \Delta X_t = \varepsilon_t \qquad (4.2)$$

Opštu klasu parametarskih modela u domenu vremena čine modeli integrisanih autoregresionih pokretnih proseka (ARIMA modeli), koji su pogodni za opisivanje stacionarnih, nestacionarnih, sezonskih i nesezonskih pojava. ([3], [5], [6], [7], [8], [13])

Definicija 17 *Operator pomeranja* B *je dat relacijom* $\Delta = 1 - B$.

Neki specijalni slučajevi ARIMA procesa:

1. ARIMA(0, 1, 0) – proces slučajnog hoda:

$$\Delta X_t = \varepsilon_t, \text{ tj. } (1-B)X_t = \varepsilon_t$$

Ako se u ovaj izraz uvede konstanta θ_0, dobija se $X_t = X_{t-1} + \theta_0 + \varepsilon_t$, pa se za X_t tada kaže da je *proces slučajnog hoda sa konstantom*.

2. ARIMA(0, 1, 1) – dat je izrazom:

$$\Delta X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1}, \text{ tj. } (1-B)X_t = (1-\theta_1 B)\varepsilon_t, \text{ gde je } |\theta_1| < 1.$$

Diferenciranjem procesa dobija se stacionarni MA(1) proces. Vrednosti autokorelacionih koeficijenata ovog procesa, kao i kod ostalih ARIMA procesa, veoma sporo opadaju ka nuli. Samo parcijalni autokorelacioni koeficijent sa kašnjenjem 1 je različit od nule i blizak jedinici.

Međutim, obična i parcijalna autokorelaciona funkcija prvih diferenci ovog procesa imaju sve karakteristike odgovarajućih funkcija kod MA(1) procesa.

3. ARIMA(0, 2, 2) – dat je izrazom:

$$\Delta^2 X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} \theta_2 \varepsilon_{t-2}, tj.(1-B^2)X_t = (1-\theta_1 B - \theta_2 B^2)\varepsilon_t$$

4. ARIMA (1, 1, 1) – dat je izrazom:

$$\Delta X_t - \phi_1 \Delta X_{t-1} = \varepsilon_t - \theta_1 \varepsilon_{t-1}, tj.(1-\phi_1 B)(1-B)X_t = (1-\theta_1 B)\varepsilon_t$$

Sadašnja vrednost procesa X_t može se izraziti preko:

i) Prethodnih vrednosti $X_s, s < t$ i sadašnje i prethodne vrednosti $\varepsilon_s, s \le t$, direktnim korišćenjem diferencne jednačine

Ako je $\varphi(B)(1-B)^d = 1 - \varphi_1 B - \varphi_2 B^2 - ... - \varphi_{p+d} B^{p+d}, \theta_0 = 0$, tada izraz za ARIMA(p,d,q) može se zapisati kao

$$X_t = \varphi_1 X_{t-1} + ... + \varphi_{p+d} X_{t-p-d} - \theta_1 \varepsilon_{t-1} - ... - \theta_q \varepsilon_{t-1} + \varepsilon_t .$$

ii) beskonačne težinske sume sadašnje i prethodnih vrednosti ε_{t-j}

$$X_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} = \psi(B)\varepsilon_t .$$

Ako se obe strane ove jednakosti pomnože sa, $\varphi(B)$, dobija se, $\varphi(B)X_t = \varphi(B)\psi(B)\varepsilon_t$, a kako je $\varphi(B)X_t = \theta(B)\varepsilon_t \Rightarrow \varphi(B)\psi(B) = \theta(B)$;

iii) beskonačne težinske sume prethodnih vrednosti $X_s, s < t$ i sadašnje vrednosti ε_t

Model $X_t = \psi(B)\varepsilon_t$ može se zapisati kao $\psi^{-1}(B)X_t = \varepsilon_t$ ili kao $\pi(B)X_t = (1 - \sum_{j=1}^{\infty} \pi_j B^j)X_t = \varepsilon.$ Dakle:

$$X_t = \pi_1 X_{t-1} + \pi_2 X_{t-2} + ... + \varepsilon_t$$

Zbog uslova invertibilnosti, $\pi(B)$ mora konvergirati ka nuli.

4.6.1 Izgradnja ARIMA modela

Opštu strategiju modeliranja ARIMA procesa koncipirali su Box[2] i Jenkins[3], pa se prema njima ona naziva i Box-Jenkinsova metoda. Da bi nastao kvalitetan model, treba da prođe kroz tri etape ove metodologije: identifikaciju modela, njegovo ocenjivanje i proveru njegove adekvatnosti i da zadovolji osnovne principe koji karakterišu dobar model ([3]):

---

[2] George Edward Pelham Box (1919 – 2013), britanski statističar.
[3] Gwilym Meirion Jenkins (1932 – 1982), britanski statističar.

1. ekonomičnost – opisati pojavu što jednostavnijim modelom, koji će istaći suštinsku karakterstiku izučavane pojave;

2. identifikabilnost – bez identifikacije modela postoje bar dva skupa vrednosti koeficijenata koji su u saglasnosti sa podacima;

3. konzistentnost sa podacima i teorijom – testovima se utvrđuje adekvatnost modela, tj. njegova usaglašenost sa podacima i apriornom znanju (ekonomskoj teoriji ili zdravom razumu).

4. prihvatljivost podataka – model ne sme da predviđa vrednosti koje ne zadovoljavaju neka ograničenja koja su prirodna za posmatranu pojavu (npr. model ne sme davati negativne vrednosti za bruto društveni proizvod);

5. uspešnost prognoziranja – kriterijum uspešnosti prognoziranja proverava se tako što se koriste opservacije van uzorka za ocenjivanje u cilju provere stepena preciznosti prognoze modela. Takođe, ako jedan od modela ima manju srednje kvadratnu grešku prognoze, a sve ostale karakteristike su jednake, tada se taj model kategoriše kao prihvatljiviji za korišćenje.

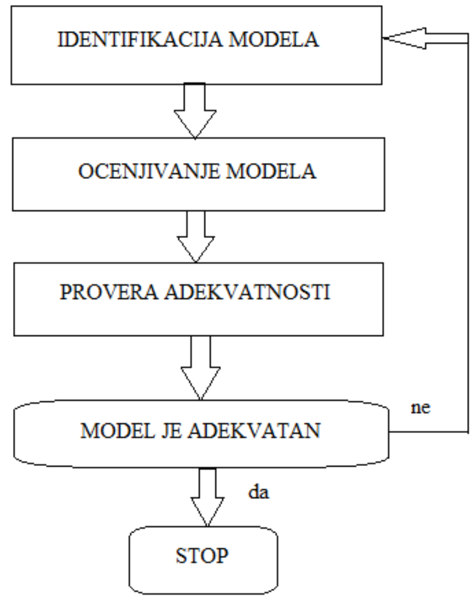
6. obuhvatnost – model treba ne samo da objasni, odnosno opiše podatke, već i da objasni uspeh ili promašaj konkurentskog modela u objašnjenju istih podataka.

Osnovu Box- Jenkins-ove metode opšte strategije modeliranja čine tri etape izgradnje modela : ([3], [5])

1. *identifikacija* - postupak korišćenja podataka vremenske serije u cilju izdvajanja uže klase ekonomičnih ARIMA modela, koji se uzimaju u razmatranje kao potencijalni generator datog skupa podataka; na osnovu grafika i korelograma najpre se utvrdi potreba za transformacijama, a zatim bira odgovarajući model;

2. *ocenjivanje* – postupak zaključivanja o koeficijentima modela na osnovu raspoloživih podataka, što je uslovljeno adekvatnošću izabranog modela; konačne ocene se dobijaju metodom najmanjih kvadrata ili metodom maksimalne verodostojnosti;

3. *provera adekvatnosti* – postupak suočavanja prilagođenog modela podacima u cilju otkrivanja njegovih eventualnih nedostataka, što podrazumeva proveru statističke značajnosti ocenjenih koeficijenata i osobina reziduala (predstavljaju li proces belog šuma); model se ili poboljšava ili se, ako zadovoljava kriterijume, koristi za prognozu.



Slika 4.1 Dijagram toka Box- Jenkinsovog iterativnog postupka [3]

4.7 Prognoziranje budućih vrednosti preko sezonskih ARIMA modela

Vremenske serije kod kojih se pojavljuju periodična kolebanja (fluktuacije) u vremenskim intervalima do godinu dana, nazivaju se sezonske vremenske serije. Često se ovakve serije sreću u turizmu, zaradama koje se isplaćuju radnicima na osnovu učinka, proizvodnji itd., čija kretanja uslovljava vreme kao doba godine (meteorološki) ili instituciona rešenja koja se primenjuju u privredi (npr. veći priliv turista u vreme praznika, povećani obim prodaje piva u letnjim mesecima, smanjenje uvoza poljoprivrednih proizvoda usled rodne godine, kolektivni godišnji odmori, itd.). U ovakvim slučajevima, izostavljanje sezonskog faktora dovelo bi do formiranja neoptimalnog modela.

Sezonski ARIMA modeli, kao i nesezonski, uzimaju u obzir međuzavisnost uzastopnih opservacija vremenske serije, npr. zavisnost opservacija za uzastopne mesece (kvartale) unutar jedne godine. Međutim, za razliku od nesezonskih, istovremeno uzimaju u obzir i međuzavisnost između opservacija za iste mesece (kvartale) u uzastopnim godinama. Ideja je dobijanje ekonomičnog ARIMA modela, koji će sa relativno malim brojem dodatnih koeficijenata u odnosu na nesezonske modele uspešno da modelira i sezonska kolebanja vremenske serije.

Najmanji vremenski period u kome se ponovi uočena pojava naziva se period sezone i označava se s. Ako su u pitanju mesečne serije, tada je period sezone jednak 12, za kvartalne serije $s=4$, a za polugodišnje $s=2$. S obzirom na to da se posmatrana pojava sa izvesnim pravilnostima ponavlja posle perioda sezone, očekuje se da će opservacije razdvojene periodom s biti međusobno korelisane. ([3])

Pored grafika sezonske vremenske serije, i korelogram sezonske vremenske serije može biti od pomoći pri utvrđivanju sezonskog karaktera serije. Lagano smanjivanje vrednosti autokorelacionih koeficijenata na sezonskim periodima (s, $2s$, $3s$,…) pokazatelj je sezonske nestacionarnosti.

Slično postupku eliminisanja nestacionarnosti kod nesezonskih serija korišćenjem operatora diferenciranja $(1-B)$, za otklanjanje sezonske nestacionarnosti koristimo operator sezonskog diferenciranja $(1-B^s)$.

Definicija 18 Sezonski ARIMA model za seriju $\{X_t, t \in T\}$ ima oblik

$$\phi(B)(1-B)^d (1-B^s)^D X_t = \theta(B)\varepsilon_t, \tag{4.3}$$

gde su operatori nesezonskog i sezonskog diferenciranja primenjeni d, odnosno D puta, redom.

Pri izgradnji sezonskog ARIMA modela primenjuje se iterativni postupak Boxa i Jenkinsa. Međutim, postoje izvesne poteškoće u primeni modela 4.3:

a) Bar jedan od polinoma $\phi(B)$ ili $\theta(B)$ mora biti minimalno reda s da bi obuhvatio autokorelaciju na sezonskim periodima. Kod, na primer, mesečnih serija to znači da red jednog od polinoma mora biti najmanje 12, pa se može desiti da će model sadržati toliki broj koeficijenata, što dalje zahteva razmatranje velikog broja modela u fazi njihovog izbora (ispitivanja značajnosti koeficijenata, upoređivanja modela sa manje i više koeficijenata itd.);

b) Interpretacija parcijalne autokorelacione funkcije (kao osnovnog sredstva za identifikaciju modela) je otežana, pa se uglavnom koristi obična autokorelaciona funkcija.

Iz navedenih razloga Box i Jenkins su pristupili definisanju klase tzv. *multiplikativnih sezonskih ARIMA modela*. ([3], [5], [13])

Pretpostavka je da je u opštem slučaju vremenska serija X_t, za koju je utvrđeno da ima karakteristike sezonske serije, modelirana korišćenjem nesezonskog ARIMA modela:

$$(1-B)^d \phi_p(B) X_t = \theta_q(B)\eta_t, \tag{4.4}$$

gde su $\phi_p = 1-\phi_1 B - \ldots - \phi_p B^p$ i $\theta_q = 1-\theta_1 B - \ldots - \theta_q B^q$ polinomi po operatoru kašnjenja, reda p i q, respektivno. Zbog sezonskog karaktera serije, proces η_t neće biti proces belog šuma. Koeficijenti autokorelacije ovog proces na sezonskim periodima biće različiti od nule. Zato se ovaj proces, takođe, modelira i ARIMA modelom

$$(1-B)^D \Phi_P(B^s)\eta_t = \Theta_Q(B^S)\varepsilon_t, \tag{4.5}$$

gde su $\Phi_P(B^S) = 1-\Phi_1 B^S - \ldots - \Phi_P B^{P_S}$ i $\Theta_Q(B^S) = 1-\Theta_1 B^S \ldots - \Theta_Q B^{Q_S}$ polinomi po operatoru B^s, reda P i Q, respektivno. Ovi polinomi zadovoljavaju standardne uslove stacionarnosti, odnosno invertibilnosti i pretpostavlja se da nemaju zajedničkih korena. Kombinujući izraze 4.4 i 4.5 dobija se *Box Jenkinsov multiplikativni sezonski ARIMA model*

$$(1-B)^d (1-B^S)^D \phi_p(B)\Phi_P(B^S) X_t = \theta_q(B)\Theta_Q(B^S)\varepsilon_t \tag{4.6}$$

Uobičajeno je da se u modelu 4.6 polinomi $\phi_p(B)$ *i* $\theta_q(B)$ *nazivaju regularnim autoregresionim polinomom i polinomom pokretnih proseka*, a $\phi_P(B^S)$ *i* $\Theta_Q(B^S)$ *sezonskim autoregresionim polinomom i polinomom pokretnog proseka*. Standardna oznaka multiplikativnih sezonskih ARIMA modela sa periodom sezone s je
ARIMA $(p,d,q)(P,D,Q)_s$.

4.7.1 ARIMA (0,1,1)(0,1,1) model

Među multiplikativnim sezonskim ARIMA modelima najčešće se koriste sledeći: $ARIMA(1,d,0)(1,D,0)_s$, $ARIMA(1,d,0)(0,D,1)_s$, $ARIMA(0,d,1)(1,D,0)_s$ i $ARIMA(0,d,1)(0,D,1)_s$, gde d i D uzimaju vrednosti 0 ili 1. ([3], [5], [6], [7])

Najčešće primenjivana klasa multiplikativnih sezonskih ARIMA modela je $ARIMA(0,1,1)(0,1,1)_s$ model oblika:

$$(1-B)(1-B^S)X_t = (1-\theta_1 B)(1-\Theta_1 B^S)\varepsilon_t,$$

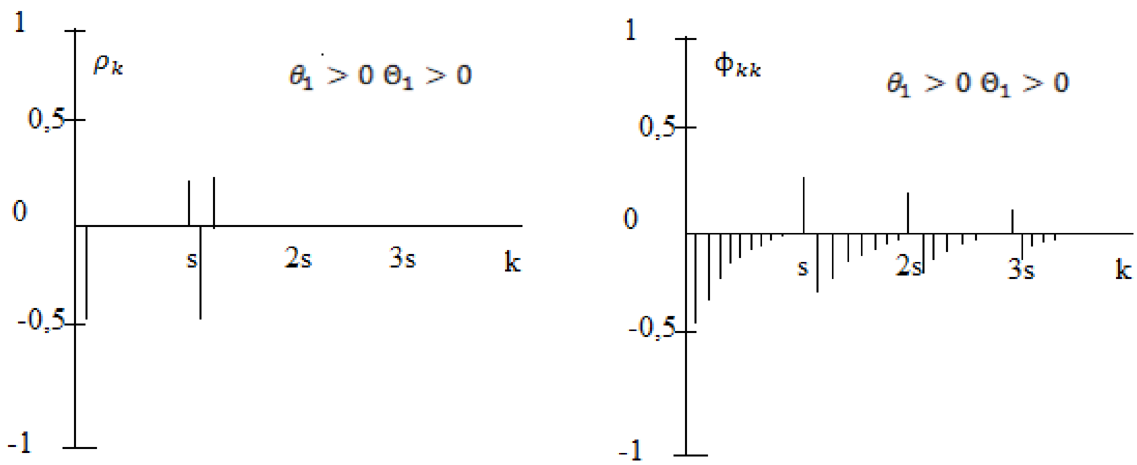
koji je u literaturi poznat kao „vazduhoplovni model”. Invertibilnost ovog modela zahteva ispunjenje uslova $|\theta_1|<1$ i $|\Theta_1|<1$.

Autokovarijaciona i autokorelaciona funkcija ovog modela su:

$$\gamma_k = \begin{cases} (1+\theta_1^2)(1+\Theta_1^2)\sigma^2, k=0 \\ -\theta_1(1+\Theta_1^2)\sigma^2, k=0 \\ \theta_1\Theta_1\sigma^2, k\in\{s-1,s+1\} \\ -\Theta_1(1+\theta_1^2)\sigma^2, k=0 \\ 0 \end{cases}$$

$$\rho_k = \begin{cases} \dfrac{-\theta_1}{1+\theta_1^2}, k=1 \\ \dfrac{\theta_1\Theta_1}{(1+\theta_1^2)(1+\theta_1^2)}, k\in\{s-1,s+1\} \\ \dfrac{-\Theta_1}{1+\Theta_1^2}, k=1 \\ 0, \end{cases}$$

Za “vazduhoplovni” model, na slici 4.2 prikazan je izgled običnog i parcijalnog korelograma za pozitivne vrednosti koeficijenta θ_1 i Θ_1.



Slika 4.2 Obična i parcijalna autokorelaciona funkcija ”vazduhoplovnog” modela [3]

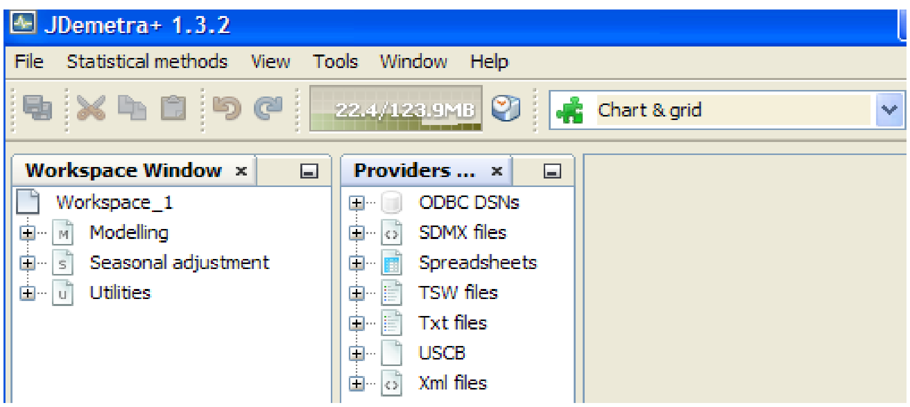
Dakle, formiranje ARIMA prognoza nije jednostavno jer, iziskuje puno računice i puno upoređivanja (u smislu koji je model adekvatnije primeniti). Prvi važan korak prilikom modeliranja jeste identifikacija modela koju zapravo čini jedan iterativni postupak koji počinje od najprostijeg ka složenijim modelima. Preliminarne ocene su bazirane na autokorelacionim koeficijentima i njihovim relacijama sa parametrima modela. Svrha preliminarnog ocenjivanja parametara je dobijanje autokorelacije odgovarajuceg modela, kao i određivanja početnih vrednosti za iterativne procedure kojima se dobijaju precizne ocene. Parcijalni autokorelacioni koeficijenti se upotrebljavaju kao sredstvo za identifikaciju autoregresionog dela ARIMA modela. Takođe se u procesu identifikacije susrećemo sa problemom reda diferencije. Kada se dobije konkretan ARIMA model, tj. model koji najbolje opisuje posmatranu seriju, tada se može pristupiti prognoziranju. Prognoza se generiše tako što uvrstimo vrednosti koje su nam poznate u predhodno dobijeni model (model je određen kad su određeni njegovi koeficijenti). U praksi se procedura identifikacije najčešće izvodi pomoću pogodnih kompjuterskih pograma najprihvatljivi je za sada JDemetra+. Pomenuti softver pored određivanja reda diferenciranja u potpunosti određuje ARIMA model što ga čini jako elegantnim i postavlja ga na prvo mesto među softverima koji služe za sezonsko prilagođavanje - desezoniranje i prognoziranje vremenskih serija. Dakle, JDemetra+ sve opisano sprovodi i dolazi do najboljeg modela – modela koji na najbolji način opisuje konkretnu seriju. Više o softreru na konkretnim primerima u glavi koja sledi.

Pored prognostičkih modela koji su opisani u ovom radu koriste se takođe i nelinearni modeli kao što su: krive rasta, logaritamske transformacije, modeli dekompozicije,…Oni se koriste u slučajevima kada serijiju nije moguće modelirati opisanim metodama i prilično su zahtevni.

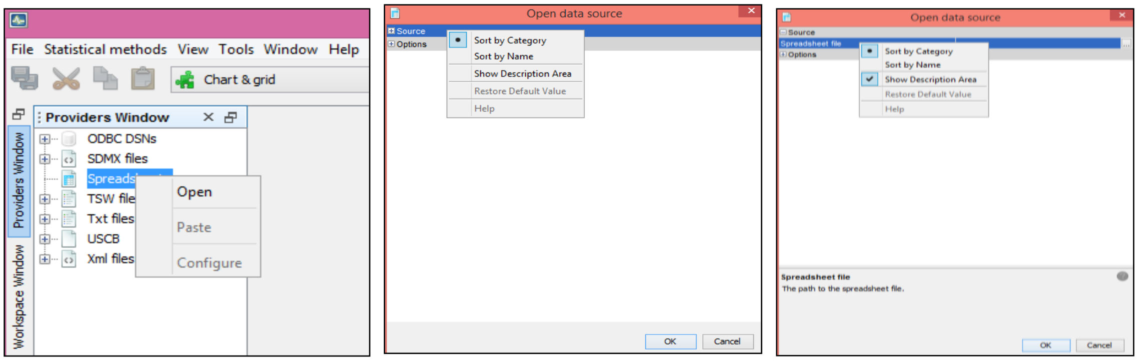
5 Prognoziranje budućih vrednosti – Uputstvo, procedura i primeri

5.0 Kratko upustvo za korištenje stoftvera JDemetra+ u svrhu prognoziranja:

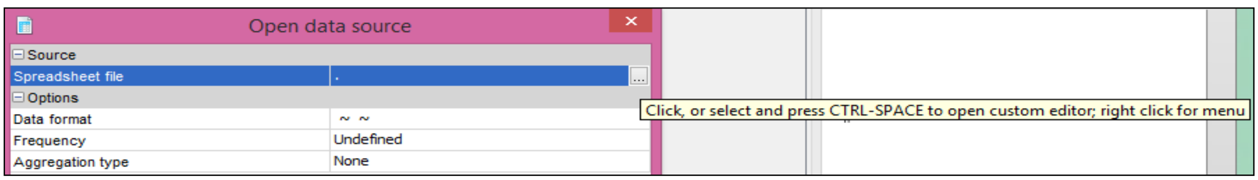
Softver Jdemetra+ je softver koji je javno dostupan svima i nalazi se na sajtu Eurostata - glavna statistička institucija Evorpske unije sa sedištem u Luksemburgu. Prvenstveno je namenjen u svrhu desezoniranja (sezonskog prilagođavanja) i prognoziranja vremenskih serija, iako se u okviru njega mogu videti mnoge druge karakteristke vremenskih serija kao i koristite razne procedure. Jednostavno se preuzima i pravolinijski se instalira na PC. Ovaj program zahteva Java podršku, Windows 2003 i više, Misrosoft Office 2003 i više.
Softver JDemetra, verzija 1.3.2 pokreće se, standardno kao i svaki drugi program na jedan od sledeća tri načina: Start/All Programs/levi taster miša na JDemetra 1.3.2 ili dvostruki levi taster miša na shortcut ikonu na desktopu ili desni taster miša na ikonu JDemetra 1.3.2/Open.
U svakom od ovih slučajeva, po difoltu sledi standardni padajući meni (File, Statistical methods, View, Tools, Window, Help), manji broj standardnih ikonica, dva aktivna prozora (Workspace Window, Providers Window) i treći prazan radni prozor. Prema potrebi prozori se aktiviraju, respektivno ili samo jedan od njih, ukoliko se radi osvežavanje podataka (Refresh). Svaki od ova tri prozora se može ručno zatvoriti ⊠, pomerati, smanjivati - (Minimize) □, povećavati (Maximize) □ - ili minimizirati (Minimize Window Group) □.



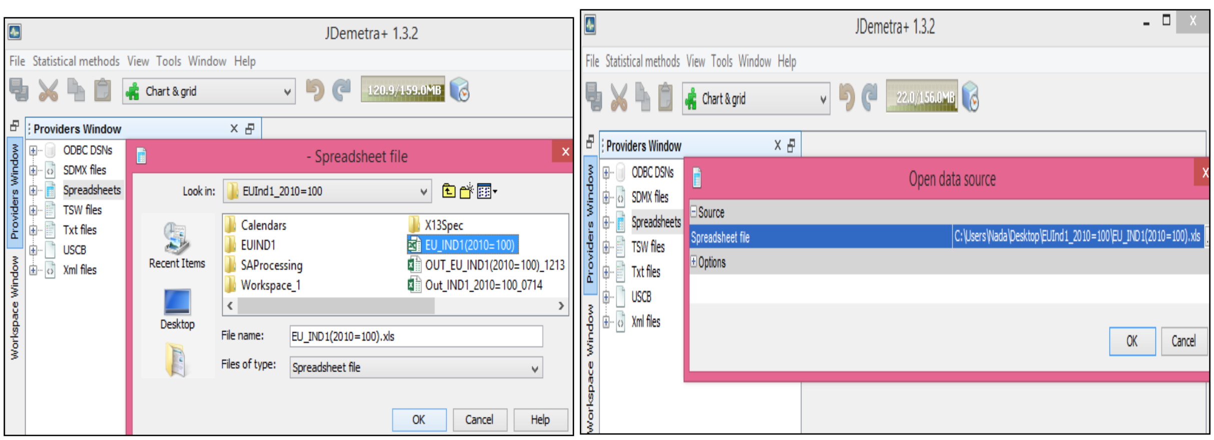
Učitavanje i forma jedne serije kao i učitavanje i forma više serija se radi aktiviranjem prozora ***Providers Window***/⊟*Spreadsheets*/*desni taster miša*/*Open*, tada se otvara prozor ***Open data sorce***. U ovom prozoru otvoriti jednu za drugom obe, veoma moćne opcije ⊞ *Source*/*levi taster miša* i ⊞ *Options*/*levi taster miša*. Postavljanjem bilo na viši nivo (⊞ *Source* ili ⊞ *Options*), bilo na niži nivo (⊟*Source*/*Spreadsheet file*) i aktiviranjem *desnog tastera miša*, otvaraju se mogućnosti sortiranja po *kategorijama* ovog prozora, sortiranja po *imenu*, kao i mogućnost *opisa* svake od njih u dnu ekrana (□*Sort by Category*, *Sort by Name*, *Show Description Area*).



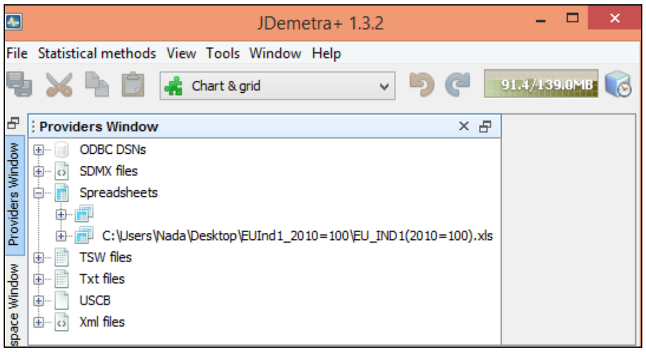
Ako se vrh strelice miša postavi na kvadratić (...)sa tri tačkice od *Spreadsheet file* i kratko zadrži dobija se poruka koja se vidi u ***Open data source*** prozoru. Ova poruka je ključ za uvoz jedne ili više serija, bilo da se to radi *levim tasterom miša* na kvadratić (...) bilo *selektovanjem/ctrl+razmaknica*, u oba slučaja sledi put za importovanje serije u program.

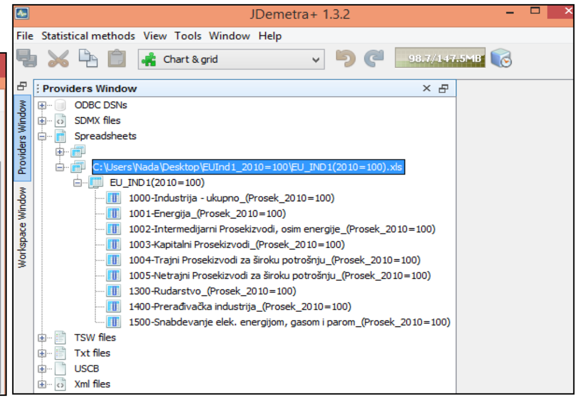


Spreadsheets file/prozor - ***Spreadsheet file****/Desktop/*EU_IND1(2010=100)/*OK*, tada u <u>prozoru</u> ***Open data source*** pod <u>kategorijom</u> *Spreadsheet file* uočava se sledeći putanja do fajla.



Treba otvoriti jedan za drugim, jedan po jedan nivo *levi taster miša/ Spreadsheets/levi taster miša /izabrati fajl koji želimo sa serijema/levi taster miša/* EU_IND1(2010=100)/*otvara se drvo sa ikonicama i spisakom serija.*





Kada su serije učitane, zatimpokrećemo proceduru: *Statistical methods/ Seasonal adjustment/ Multi Processing /New*. Zatim serije "prevučemo" u prozor *Drop here* koji nam se otvorio sa desne strane. Sva podešavanja ostavimo na *default* vrednost i pritiskom na zelenu strelicu dobijamo prognoze. U pozadini programa se nalaze sve procedure ranije opisane koje su se automatski sprovele, tako je dobijen adekvatam ARIMA model i izračunate su prognoze. Prognoze čitamo tako što idemo u donji levi ugao na *Main results/table*. U tabeli su vrednosti predstavljene ***italic*** fontom su zapravo prognoze. Na ovom mestu dovoljno nam je ovoliko poznavanje rukovanja softverom. Više o korištenju softvera u [12].

Kako praktično prognozirati serije?

5.1 Procedura:

A) Priprema serije:

A1 Crtanje grafikona

A2 Donošenje odluke o momentu od kojeg je adekvatno posmatrati seriju

A3 "Presecanje" serije na adekvatnim mestima

A4 Spremanje fajla za ulaz u softver JDemetru+

B) Primena ARIMA modela za prognoziranje

C) Ekspertska korekcija prognoze dobijene ARIMA modelom – koja se formira polazeći od modelske prognoze, posmatranjem ***periodičnih stopa rasta*** i subjektivnom korekcijom zasnovanom na iregularnosti u prošlosti (koliko će po mišljenju eksperta da se nepredviđena situacija iz prošlosti odraziti na pojavu koju pratimo). Kao i koliko će se predviđena neregularnost u budućnosti uticati na tok serije.

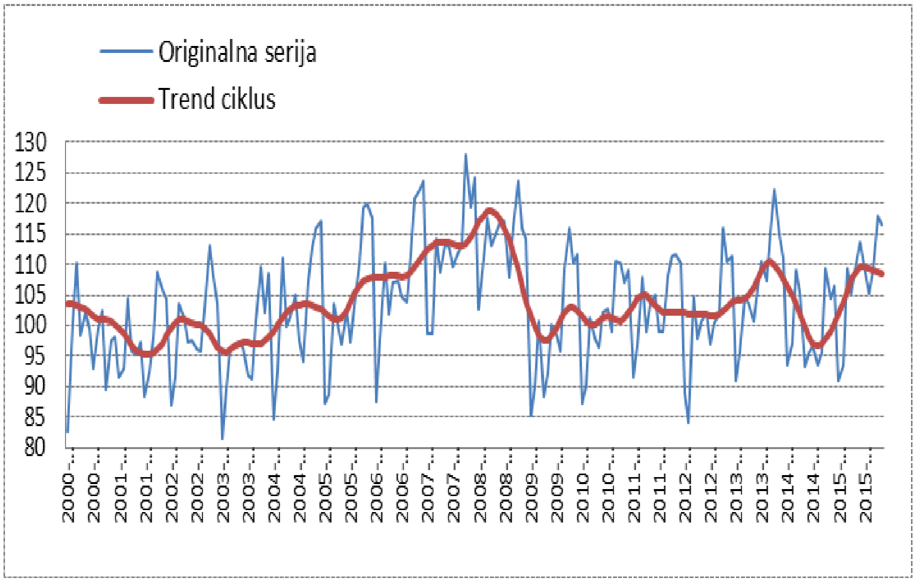
5.2 Primeri

5.2.1 Prognoziranje vremenske serije INDUSTRIJSKE PROIZVODNJE:

A Priprema serije:

A1 Crtanje grafika:

Na slici 5.1 prikazani su bazni indeksi industrijske proizvodnje, za period od januara 2000. godine do novembra 2015. godine. Najjednostavnije je nacrtati grafikon koristeći EXCEL, mada je to moguće uraditi i u drugom programu. Razlog više da se radi koristeći EXCEL je taj da se podaci dobijeni iz baze vremenskih serija Republičkog zavoda za statistiku ([11]) već nalaze u .xls formatu.



Slika 5.1 Mesečni indeksi industrijske proizvodnje u Republici Srbiji, od 2000. do 2015. godine (2014=100)

A2 Donošenje odluke o momentu od kojeg je adekvatno posmatrati seriju:

Kao što se može videti sa grafika, do 2008. god. serija je imala stabilan rastući trend, ali sa dolaskom ekonomske krize, koja je nastupila pomenute godine, serija beleži značajan pad koji se naziva ***strukturni lom***. Od tog momenta serija nastavlja da "živi" na nivou dosta nižem od prethodnog (nema bitnijih iregularnih promena).

A3 "Presecanje" serije na adekvatnim mestima:

Očigledno, pogodno mesto za "presecanje" serije je upravo 2008. godina, što znači da ćemo za našu prognozu koristiti seriju od 2009. pa do 2015. godine. U procesu dobijanja najpouzdanije prognoze, prvo ćemo uporediti nekoliko različitih raspona jedne iste serije, ne bi li smo dobili najpribližniju prognozu modela. Pod tim se konkretno podrazumeva da uzmemo npr. *šest* vremenskih raspona serije (ukoliko želimo da imamo šest modela), tj. da seriju "isečemo" na šest pogodnih različitih mesta, vodeći računa da serija ima aproksimativno ceo broj godina. Pošto je u našem slučaju poslednji realizovani indeks u novembru 2015. godine, i treba da prognoziramo decembar, januar i februar, napravićemo 6 serija koje počinju različitim datumima, a završavaju se novembrom 2015. godine. To znači sledeće:

- u prvom modelu uzećemo seriju raspona XI 2009 - XI 2015;
- u drugom XII 2009 - XI 2015.;
- u trećem XI 2010 - XI 2015.;
- u četvrtom XII 2010 - XI 2015.;
- u petom XI 2011 - XI 2015.
- i u šestom XII 2011 - XI 2015. godine.

A4 Spremanje fajla za ulaz u JDemetru+:

Ovaj deo je detaljno opisan u Upustvu za korištenje softvera JDemetra+ ([12]). Ukratko, eksel fajl u kojem se nalaze serije (različiti rasponi iste serije) treba srediti tako da prva kolona, u kojoj su datumi, bude u određenom formatu (datum oblika "jan-09" levo poravnata).

B. Primena ARIMA modela za prognoziranje

Šest modela serija iz tačke A3 (koje se nalaze u eksel fajlu) učitavamo u JDemetra+ softver i pokrećemo proceduru *Statistical methods => Seasonal adjustment => Multi Processing => New*. U specifikaciji ćemo ostaviti *default* podešavanja. Zatim pokrećemo *multi processing*, takođe po default podešavanju JDemetra+. Bitno je adekvatno izabrati tip ARIMA modela - njegove parametre, našim serijama se bira sezonski ARIMA model iz razloga što serije imaju izraženu sezonsku komponentu. ARIMA modeliranje nije ni malo lak proces (računanje autokorelacione funkcije, parcijalne autokorelacione funkcije, određivanje broja značajnih koeficijenata u AR i MA modelima itd. što se može videti iz ranije predstavljenog ARIMA modela), ali na našu sreću - to sve odradi softver JDemetra+ umesto nas. Pošto smo imali više modela (više dužina jedne te iste serije), imaćemo i više prognoza (u ovom primeru šest). Opšte pravilo glasi: **odbaciti najmanju i najveću** vrednost. Od ***preostale 4 vrednosti naći ćemo prosek i dobijene vrednosti biće tražene ARIMA prognoze.***
Izlazne tabele posle obrade iz JDemetra+ izgledaju ovako:

	Series
3-2014	108.9
4-2014	105.4
5-2014	93.3
6-2014	95.6
7-2014	96.8
8-2014	93.5
9-2014	95.6
10-2014	109.4
11-2014	104.4
12-2014	106.6
1-2015	91
2-2015	93.7
3-2015	109.3
4-2015	105.4
5-2015	110.1
6-2015	113.7
7-2015	110
8-2015	105.3
9-2015	109
10-2015	117.9
11-2015	116.6
12-2015	*114.27*
1-2016	*93.444*
2-2016	*94.807*

Model 1: XI 2009-XI 2015

	Series
3-2014	108.9
4-2014	105.4
5-2014	93.3
6-2014	95.6
7-2014	96.8
8-2014	93.5
9-2014	95.6
10-2014	109.4
11-2014	104.4
12-2014	106.6
1-2015	91
2-2015	93.7
3-2015	109.3
4-2015	105.4
5-2015	110.1
6-2015	113.7
7-2015	110
8-2015	105.3
9-2015	109
10-2015	117.9
11-2015	116.6
12-2015	*116.811*
1-2016	*98.015*
2-2016	*100.948*

Model 2: XII 2009-XI 2015

	Series
3-2014	108.9
4-2014	105.4
5-2014	93.3
6-2014	95.6
7-2014	96.8
8-2014	93.5
9-2014	95.6
10-2014	109.4
11-2014	104.4
12-2014	106.6
1-2015	91
2-2015	93.7
3-2015	109.3
4-2015	105.4
5-2015	110.1
6-2015	113.7
7-2015	110
8-2015	105.3
9-2015	109
10-2015	117.9
11-2015	116.6
12-2015	*115.551*
1-2016	*97.149*
2-2016	*99.359*

Model 3: XI 2010-XI 2015

	Series
3-2014	108.9
4-2014	105.4
5-2014	93.3
6-2014	95.6
7-2014	96.8
8-2014	93.5
9-2014	95.6
10-2014	109.4
11-2014	104.4
12-2014	[illegible]
1-2015	91
2-2015	93.7
3-2015	109.3
4-2015	105.4
5-2015	110.1
6-2015	113.7
7-2015	110
8-2015	105.3
9-2015	109
10-2015	117.9
11-2015	116.6
12-2015	115.917
1-2016	98.226
2-2016	101.149

Model 4: XII 2010-XI 2015

	Series
3-2014	108.9
4-2014	105.4
5-2014	93.3
6-2014	95.6
7-2014	96.8
8-2014	93.5
9-2014	95.6
10-2014	109.4
11-2014	104.4
12-2014	[illegible]
1-2015	91
2-2015	93.7
3-2015	109.3
4-2015	105.4
5-2015	110.1
6-2015	113.7
7-2015	110
8-2015	105.3
9-2015	109
10-2015	117.9
11-2015	116.6
12-2015	114.705
1-2016	96.681
2-2016	98.07

Model 5: XI 2011-XI 2015

	Series
3-2014	108.9
4-2014	105.4
5-2014	93.3
6-2014	95.6
7-2014	96.8
8-2014	93.5
9-2014	95.6
10-2014	109.4
11-2014	104.4
12-2014	106.6
1-2015	91
2-2015	93.7
3-2015	109.3
4-2015	105.4
5-2015	110.1
6-2015	113.7
7-2015	110
8-2015	105.3
9-2015	109
10-2015	117.9
11-2015	116.6
12-2015	115.6
1-2016	96.965
2-2016	98.439

Model 6: XII 2011-XI 2015

Gledajući u svim modelima ***decembar 2015.*** primećujemo da je najmanja vrednost u modelu 1 (114,27), a najveća u modelu 2 (116,81), pa prema opisanoj proceduri, te dve vrednosti eliminišemo iz razmatranja. Postmatrajući sada u svim modelim ***januar 2016.*** primećujemo da je najmanja vrednost u modelu 1 (93,44), a najveća u modelu 4 (98,23), pa po opisanoj proceduri te dve vrednosti eliminišemo iz razmatranja. Na isti način postupamo kad je u pitanju ***februar 2016.***

C Ekspertska korekcija prognoze dobijene ARIMA modelom:

Kako ARIMA jednako radi u Americi, Evropi i Australiji (na svim podacima na isti način prognozira – sprovodi ranije opisanu proceduru) potrebna je korekcija eksperta sa velikim iskustvom, ne samo u oblasti vremenskih serija već i na raznim drugim poljima kao npr. kretanju raznih pokazatelja koji su direktno utiču na pojavu – seriju koju prognoziramo. Ta korekcija podrazumeva ***uvećanje ili umanjenje vrednosti dobijenih ARIMA*** modelom, na osnovu dešavanja u zemlji i inostranstvu. Dešavanja koja na to utiču su: poplave, zemljotresi, ratovi, migranti, razne odluke Vlade npr. povećanje ili smanjenje akcize, i pojave slične ovima, koje su se desile u prošlosti, a mogle bi se reflektovati na seriju koju prognoziramo. Konkretno, u primeru industrijske proizvodnje, ekspertska procena se vrši zbog poplave u maju 2014. godine, koja je pogodila energetski sektor industrije, a koji se u potpunosti oporavio sredinom 2015. godine, tako da dobijene prognoze korigujemo ‘malo naviše’ i to ćemo raditi do sredine godine, osim u slučaju da se u međuvremenu ne dogodi nešto nepredviđeno.
U Tabeli 5.1 su prikazane prognoze (prosečne vrednosti dobijene od 4 preostale, nakon eliminisanja najviše i najniže) koje daje ARIMA model.
U Tabeli 5.2 su prikazane finalne prognoze – koje sadrže eksperstku korekciju.

Tabela 5.1– Izlaz iz JDemetra+ (aritmetička sr.)

Год.	**I** *Jan*	**II** *Feb*	**III** *Mar*	**IV** *Apr*	**V** *May*	**VI** *Jun*	**VII** *Jul*	**VIII** *Aug*	**IX** *Sep*	**X** *Oct*	**XI** *Nov*	**XII** *Dec*
2013	91,1	95,6	104,9	103,7	100,6	106,1	110,3	107,3	115,4	122,1	114,8	111,1
2014	93,6	96,9	108,9	105,4	93,3	95,6	96,8	93,5	95,6	109,4	104,4	106,6
2015	91,0	93,7	109,3	105,4	110,1	113,7	110,0	105,3	109,0	117,9	116,6	***115,4***
2016	***97,2***	***99,2***										

Tabela 5.2– Konačna prognoza

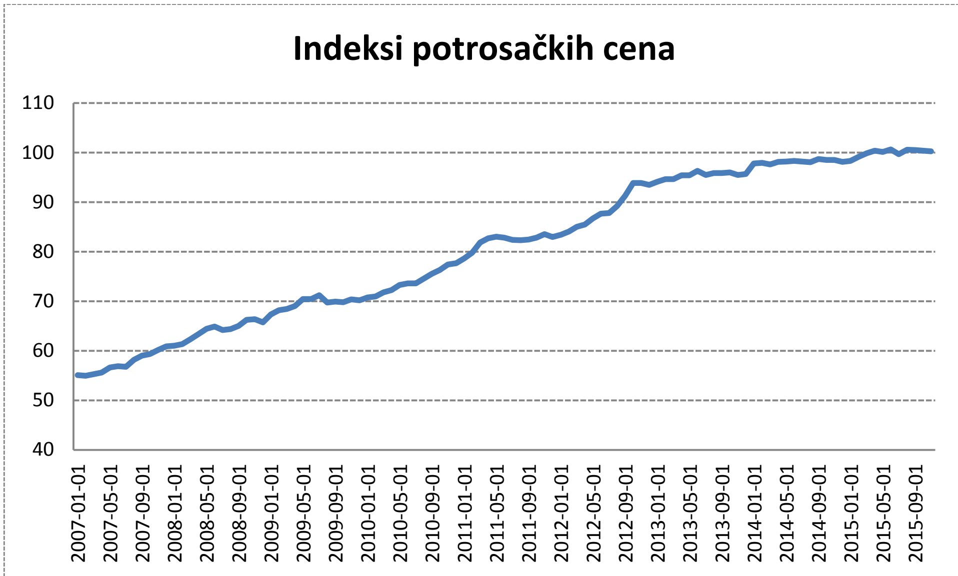
Год.	**I** *Jan*	**II** *Feb*	**III** *Mar*	**IV** *Apr*	**V** *May*	**VI** *Jun*	**VII** *Jul*	**VIII** *Aug*	**IX** *Sep*	**X** *Oct*	**XI** *Nov*	**XII** *Dec*
2013	91,1	95,6	104,9	103,7	100,6	106,1	110,3	107,3	115,4	122,1	114,8	111,1
2014	93,6	96,9	108,9	105,4	93,3	95,6	96,8	93,5	95,6	109,4	104,4	106,6
2015	91,0	93,7	109,3	105,4	110,1	113,7	110,0	105,3	109,0	117,9	116,6	***116,3***
2016	***97,5***	***100,4***										

Neke <u>bitne</u> preporuke:

1. Serija NE sme biti kraća od 4-5 godina.
2. Finalna prognoza se definiše kao ceo broj, eventualno kao pola broja (npr. nema baš smisla reći "prognoziram da će vrednost u decembru biti 114.83 - u ovom slučaju rekli bismo da će vrednost biti oko 114.5 ili 115, tačnije bi bilo reći oko 115, jer je bliža toj vrednosti).
3. O kvalitetu prognoze možemo diskutovati u momentu kada bude poznat originalni podatak za period koji smo prognozirali. Naime, prognoza je dobra ako se od originalne vrednosti ne razlikuje za više od 2 indeksna poena (u nekim literaturama čak i do 3 indeksna poena). Podrazumeva se da što je manja razlika između originalnog i prognoziranog podatka, to je prognoza bila bolja.
4. Prognoze se daju kao kumulativna stopa izražena u procentima. To konkretno znači da se izračunava odnos, u čijem brojiocu se nalazi suma indeksa istih meseci prethodne godine dok je u imeniocu suma indeksa istih meseci tekuće godine, pa se ovako dobijeni odnos pomonoži sa 100. Dakle, npr. ako imamo podatke za januar i februar npr 2016. godine, a prognoziramo mart, april i maj, naša prognoza će zapravo biti samo jedan podatak (broj u procentima - stopa) koji ćemo dobiti tako što sumu indeksa prvih per meseci 2015. godine, podelimo sa sumom koju čine originalni podaci za januar i februar i naše procene za mart, april i maj, pa ćemo tako dobijeni odnos pomnožiti sa 100 i to će biti prognoza "period na period" – period 2016. god u odnosu na isti period prethodne godine. Odavde se može zaključiti da je zapravo najteže prognozirati prvih par meseci, čak i do realizacije podatka za jun, a kako se bliži kraj godine manja je verovatnoća greške. To je posledica toga što krajem godine naše procene dosta manje doprinose sumi - odnosu, nego sa početka kada zapravo samo one i čine sumu.

5.2.2 Prognoziranje vremenske serije POTROŠAČKIH CENA

Poštujući proceduru Box-Jenkinsa, prvo crtamo grafik.



Slika 5.2 Indeks potrošačkih cena, mesečni indeksi, 2014=100

Sa grafikona se vidi da je serija „glatka“, sa blago rastućim trendom i sa tendencijom usporavanja rasta poslednjih godina. U cilju ocene parametara ARIMA modela seriju „sečemo“, zbog preporuke da dužina serije bude minimum četiri godine. S obzirom da je serija cena glatka (nema strukturni lom) i da se po tome razlikuje od serije Industrije, možemo je seći na više mesta, tj. ne postoji neka prelomna tačka koja bi jedina bila validna u procesu određivanja početka serije. Ovu seriju možemo prognozirati i „seći“ od 2007. godine pa nadalje do 2010. godine, imajući u vidu da serija obuhvati najmanje četiri godine i vodeći računa da u prvoj varijanti serija počinje od početka godine, a u drugoj da svaki mesec bude zastupljen u modelu isti broj puta. Posle pripreme fajla za ARIMA modeliranje u JDemetri+ pokrećemo proceduru *Statistical methods => Seasonal adjustment => Multi Processing => New,* kao što je opisano u Tački B, i sledimo dalja uputstva i korake koji su već opisani na primeru Industrije.

Tabela 5.3 – Izlaz iz JDemetra+ (aritmetička sr.)

Год.	**I** *Jan*	**II** *Feb*	**III** *Mar*	**IV** *Apr*	**V** *May*	**VI** *Jun*	**VII** *Jul*	**VIII** *Aug*	**IX** *Sep*	**X** *Oct*	**XI** *Nov*	**XII** *Dec*
2013	94,1	94,6	94,6	95,4	95,4	96,4	95,5	95,9	95,9	96,0	95,5	95,7
2014	97,8	97,9	97,6	98,2	98,2	98,4	98,2	98,0	98,7	98,5	98,5	98,1
2015	98,3	99,2	99,9	100,4	100,1	100,6	99,7	100,6	100,5	100,4	100,2	***102,4***
2016	***102,8***	***103,4***										

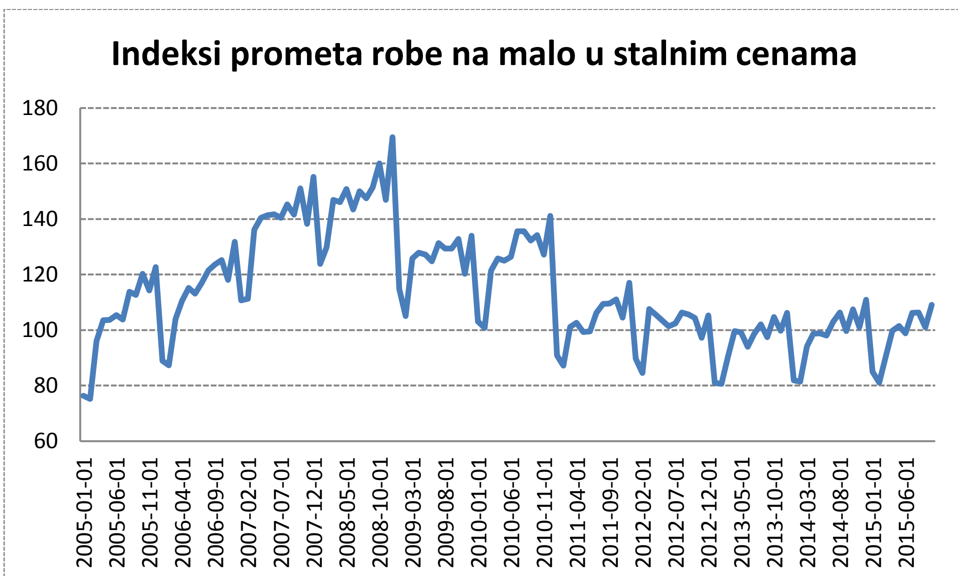
Zatim sledi takozvana ekspertska korekcija koja se zasniva na poznavanju serija cena. Npr, ako su odlukom Vlade povećane cene cigareta za 10% (ili se planira povećanje), prognozu korigujemo naviše na osnovu učešća grupe ’’duvan’’ u formiranju ukupnog indeksa potrošačkih cena koje iznosi 4%, tako da se prognoza koriguje za 0,4% naviše. Po istom principu, ako se cene naftinih derivata smanje za 5%, onda se prognoza koriguje naniže za oko 0,3%, imajući u vidu učešće grupe „naftini derivati“ u formiranju ukupnog indeksa potrošačkih cena koje je oko 6%. (Pod pretpostavkom da se cene nekog proizvoda povećaju 100% onda bi rast ukupnog indeksa potrošačkih cena bilo procentualno učešće tog proizvoda u formiranju ukupnog indeksa). U ranijim godinama (2013, 2014) bio je izražen rast cena u decembru i januaru, zbog povećanja PDV-a, a nekih drugih godina zbog povećanja cena cigareta i naftinih derivata, dok ove godine takvih najava nije bilo. Kako se autokorelaciona (ACF) i parcijalna autokorelaciona funkcija (PACF) ocenjuju na osnovu podataka iz prošlosti (tj. ACF i PACF „pamte“ podatke iz prošlosti), ARIMA model će preceniti progonoze za ove mesece, pa ih zbog toga korigujemo naniže.

Tabela 5.4 – Konačna prognoza

Год.	**I** *Jan*	**II** *Feb*	**III** *Mar*	**IV** *Apr*	**V** *May*	**VI** *Jun*	**VII** *Jul*	**VIII** *Aug*	**IX** *Sep*	**X** *Oct*	**XI** *Nov*	**XII** *Dec*
2013	94,1	94,6	94,6	95,4	95,4	96,4	95,5	95,9	95,9	96,0	95,5	95,7
2014	97,8	97,9	97,6	98,2	98,2	98,4	98,2	98,0	98,7	98,5	98,5	98,1
2015	98,3	99,2	99,9	100,4	100,1	100,6	99,7	100,6	100,5	100,4	100,2	***100,1***
2016	***100,7***	***100,8***										

5.2.3 Prognoziranje vremenske serije PROMET U STALNIM CENAMA

Poštujući proceduru Box-Jenkinsa, prvo crtamo grafik.



Slika 5.3 Indeks prometa na malo u stalnim cenama, mesečni indeksi, 2014=100

Sa grafika se vidi da je serija do kraja 2008. godine ima tendenciju rasta, da bi od početka 2009. do kraja 2010. imala blagi pad, koji se od 2011. nastavio, da bi se kasnije serija stabilizovala na osetno nižem nivou nego 2008. godine.

Zbog ocene parametara ARIMA modela seriju ’’sečemo’’ od januara 2011. godine, jer se tada serija stabilizovala i nema velikih lomova, kao što se vidi sa grafikona. Kao i u prethodnom primeru, pripremićemo fajl za ARIMA modeliranje u JDemetri+, pokrenuti proceduru *Statistical methods => Seasonal adjustment => Multi Processing => New,* kao što je opisano u Tački B, i slediti dalja uputstva i korake koji su već opisani na primeru Industrije. Ekstremne vrednosti prognoziranih vrednosti odbacimo i izaberemo ’’prognozu’’ kao aritmetičku sredinu ostalih prognoza (tabela 5.5).

Tabela 5.5 - Izlaz iz JDemetra+ (aritmetička sr.)

Год.	**I** *Jan*	**II** *Feb*	**III** *Mar*	**IV** *Apr*	**V** *May*	**VI** *Jun*	**VII** *Jul*	**VIII** *Aug*	**IX** *Sep*	**X** *Oct*	**XI** *Nov*	**XII** *Dec*
2013	81,0	80,6	93,0	99,7	99,1	94,0	98,6	102,0	97,5	104,7	99,7	106,2
2014	81,9	81,4	94,1	98,8	98,8	98,0	103,0	106,3	99,6	107,5	100,1	110,9
2015	85,0	81,0	90,3	99,8	101,5	98,8	106,3	106,4	101,0	109,1	101,8	***114,9***
2016	***88,3***	***83,3***										

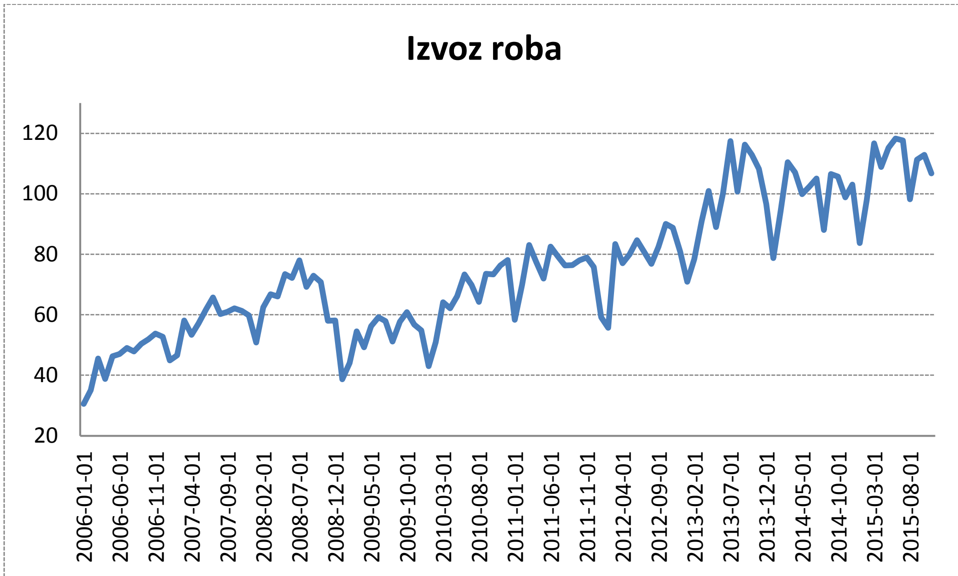
Zatim sledi takozvana ekspertska prognoza ili prognoza autora, koja se zasniva na dodatnim informacijama o kretanju pojave. Kod prometa na malo u stalnim cenama vrlo važno je pratiti kako rast ***zarada*** (promet robe na malo u tekućim cenama je u veoma visokoj pozitivnoj korelaciji sa zaradama) tako i kretanje ***cena***, jer se indeksi prometa robe u stalnim cenama dobijaju deflacioniranjem indeksa prometa robe u trgovini na malo indeksom potrošačkih cena **grupe ’’roba’’**. Na osnovu očekivanih kretanja zarada (naniže) i cena (takođe naniže) vrši se korekcija naniže prognoza dobijenih ARIMA modelom (tabela 5.6).

Tabela 5.6 – Konačna prognoza

Год.	**I** *Jan*	**II** *Feb*	**III** *Mar*	**IV** *Apr*	**V** *May*	**VI** *Jun*	**VII** *Jul*	**VIII** *Aug*	**IX** *Sep*	**X** *Oct*	**XI** *Nov*	**XII** *Dec*
2013	81,0	80,6	93,0	99,7	99,1	94,0	98,6	102,0	97,5	104,7	99,7	106,2
2014	81,9	81,4	94,1	98,8	98,8	98,0	103,0	106,3	99,6	107,5	100,1	110,9
2015	85,0	81,0	90,3	99,8	101,5	98,8	106,3	106,4	101,0	109,1	101,8	***113,8***
2016	***87,5***	***82,6***										

5.2.4 Prognoziranje vremenske serije IZVOZ ROBA

Poštujući proceduru Box-Jenkinsa, prvo crtamo grafik.



Slika 5.4 Indeks izvoza roba u evrima, mesečni indeksi, 2014=100

Na osnovu grafika vidimo da serija do kraja 2008. godine ima tendenciju rasta, da bi od početka 2009. do kraja 2010. imala blagi pad, koji se od 2011. nastavio i kasnije se serija stabilizovala. Zbog ocene parametara ARIMA modela seriju ''sečemo'' od januara 2011. godine, jer se tada serija stabilizovala i nema velikih lomova, kao što se vidi sa grafikona. Kao i u prethodnom primeru, pripremićemo fajl za ARIMA modeliranje u JDemetri+, pokrenuti proceduru *Statistical methods => Seasonal adjustment => Multi Processing => New*, kao što je opisano u Tački B, i slediti dalja uputstva i korake koji su već opisani na primeru Industrije. Ekstremne vrednosti prognoziranih vrednosti odbacimo i izaberemo ''prognozu'' kao aritmetičku sredinu ostalih prognoza (tabela 5.7).

Tabela 5.7 Izlaz iz JDemetra+ (aritmetička sr.)

Год.	**I** *Jan*	**II** *Feb*	**III** *Mar*	**IV** *Apr*	**V** *May*	**VI** *Jun*	**VII** *Jul*	**VIII** *Aug*	**IX** *Sep*	**X** *Oct*	**XI** *Nov*	**XII** *Dec*
2013	70,9	78,4	90,9	101,0	89,0	100,3	117,5	100,7	116,3	112,9	108,2	96,6
2014	78,8	94,9	110,4	107,0	99,9	102,4	105,0	88,1	106,5	105,7	98,8	103,0
2015	83,7	98,0	116,7	108,8	115,2	118,3	117,6	98,2	111,3	112,9	106,7	***106,9***
2016	***86,0***	***96,5***										

Prognozu korigujemo na osnovu kretanja industrijske i poljoprivredne proizvodnje, sa kojima je izvoz u pozitivnoj korelaciji. Ako proizvodnja raste, očekuje se i rast izvoza i obratno. Serija je u decembru korigovana naviše, jer se zbog boljeg poslovne godine očekuje da izvoz premaši „modelsku“ vrednost.

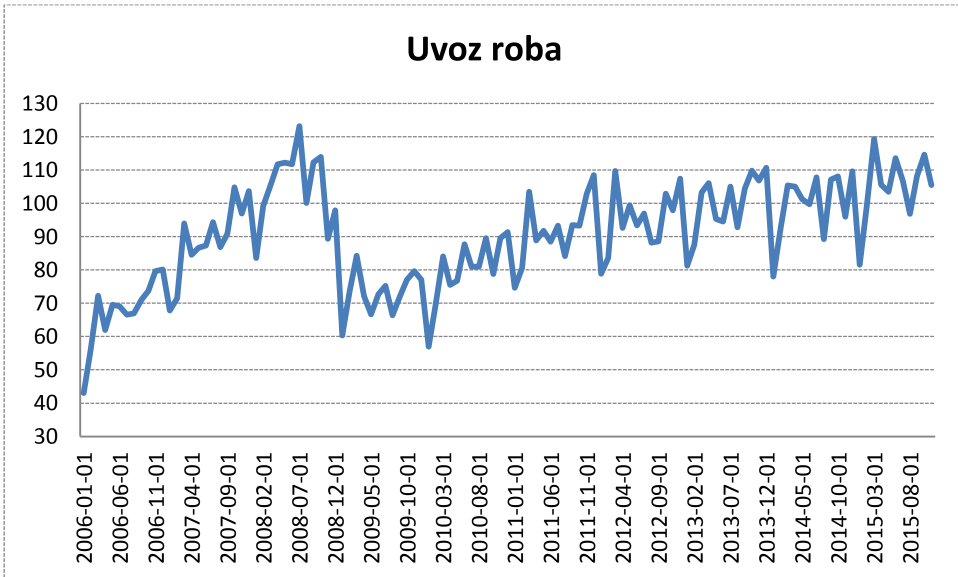
Zbog očekivanog rasta industrijske proizvodnje u prvom kvartalu prognoze su korigovane naviše i za januar i februar 2016. godine.

Tabela 5.8 – Konačna prognoza

Год.	**I** *Jan*	**II** *Feb*	**III** *Mar*	**IV** *Apr*	**V** *May*	**VI** *Jun*	**VII** *Jul*	**VIII** *Aug*	**IX** *Sep*	**X** *Oct*	**XI** *Nov*	**XII** *Dec*
2013	70,9	78,4	90,9	101,0	89,0	100,3	117,5	100,7	116,3	112,9	108,2	96,6
2014	78,8	94,9	110,4	107,0	99,9	102,4	105,0	88,1	106,5	105,7	98,8	103,0
2015	83,7	98,0	116,7	108,8	115,2	118,3	117,6	98,2	111,3	112,9	106,7	***108,9***
2016	***86,3***	***101,9***										

5.2.5 Prognoziranje vremenske serije UVOZA ROBA

Poštujući proceduru Box-Jenkinsa, prvo crtamo grafik.



Slika 5.5 Indeks uvoza roba u evrima, mesečni indeksi, 2014=100

Na osnovu grafika vidimo da serija do kraja 2008. godine ima tendenciju rasta, da bi od početka 2009. do kraja 2010. imala blagi pad, koji se od 2011. nastavio i kasnije se serija stabilizovala na na osetno nižem nivou nego 2008. godine. Zbog ocene parametara ARIMA modela seriju ’’sečemo’’ od 2009. godine, jer se tada serija stabilizovala i nema velikih lomova, kao što se vidi sa grafika. Kao i u prethodnom primeru, pripremićemo fajl za ARIMA modeliranje u JDemetri+, pokrenuti proceduru *Statistical methods => Seasonal adjustment => Multi Processing => New*, kao što je opisano u Tački B, i slediti dalja uputstva i korake koji su već opisani na primeru Industrije. Ekstremne vrednosti prognoziranih vrednosti odbacimo i izaberemo ’’prognozu’’ kao aritmetičku sredinu ostalih prognoza (tabela 5.9).

Tabela 5.9 - Izlaz iz JDemetra+ (aritmetička sr.)

Год.	**I** *Jan*	**II** *Feb*	**III** *Mar*	**IV** *Apr*	**V** *May*	**VI** *Jun*	**VII** *Jul*	**VIII** *Aug*	**IX** *Sep*	**X** *Oct*	**XI** *Nov*	**XII** *Dec*
2013	81,3	87,5	103,3	106,1	95,3	94,6	105,0	92,8	104,5	109,9	106,9	110,7
2014	78,0	92,7	105,4	105,0	101,3	99,7	107,8	89,2	107,2	108,1	96,0	109,6
2015	81,6	99,1	119,3	105,6	103,5	113,5	106,7	96,9	108,2	114,6	105,5	***117,9***
2016	***82,8***	***99,2***										

Na osnovu dodatnih informacija o uvozu roba (serija je u visokoj korelaciji sa industrijskom i poljoprivrednom proizvodnjom, jer ako podbaci poljoprivredna proizvodnja nekih proizvoda mesto da izvozimo te proizvode mi ih uvozimo) korigujemo prognozu. Serija uvoza roba je nezahvalna za ekspertsku korekciju, jer zavisi od likvidnosti uvoznika tako da je teško predvideti u kom mesecu će recimo biti uvoz energenata ili uvoz repromaterijala i robe široke potrošnje. Konkretno, u decembru obično dolazi do utvrđivanja poslovnog rezultata, a u cilju njegovog boljeg prikazivanja smanjuje se uvoz, koji se kasnije – u januaru i februaru povećava da bi kompenzovao ovaj pad. Zato decembarsku prognozu korigujemo naniže, a januarsku i februarsku naviše.

Tabela 5.10 – Konačna prognoza

Год.	**I** *Jan*	**II** *Feb*	**III** *Mar*	**IV** *Apr*	**V** *May*	**VI** *Jun*	**VII** *Jul*	**VIII** *Aug*	**IX** *Sep*	**X** *Oct*	**XI** *Nov*	**XII** *Dec*
2013	81,3	87,5	103,3	106,1	95,3	94,6	105,0	92,8	104,5	109,9	106,9	110,7
2014	78,0	92,7	105,4	105,0	101,3	99,7	107,8	89,2	107,2	108,1	96,0	109,6
2015	81,6	99,1	119,3	105,6	103,5	113,5	106,7	96,9	108,2	114,6	105,5	***116.1***
2016	***83,4***	***103.9***										

6 ZAKLJUČAK

U ovom radu prikazana je teorija koja opisuje AR, MA, ARMA, eksponencijalno izravnanje, ARIMA i sezonski ARIMA modele. Oni imaju za cilj da modeliraju vremensku seriju i da na osnovu podataka dostupnih do sada prognoziraju buduće vrednosti. Ovi modeli su jako rasprostanjeni u savremenoj ekonometriji. Daleko najrasprostranjeniji model je svakako sezonski ARIMA model, jer je univerzalan, tj. adekvatan za modeliranje stacionarnih i nestacionarnih serija koje uz to imaju izraženu sezonsku komponentu.

Neki od zadataka za budući rad na ovom polju je pravljenje modela (modeliranje) koji prepoznaju značajna odstupanja u vremenskoj seriji i pronalaženje načina da se to odstupanje uključi u prognozu.

"*Svi modeli su pogrešni, ali neki su sigurno korisni*" rekao je *George E.P. Box*. Modeliranje je, kao što se može zaključiti, jako kompleksan proces iz jednostavnog razloga što se mahom oslanja na intuiciju i iskustvo autora sa velikom primesom subjektivnosti. Na procene se treba oslanjati, samo ako je problem dobro definisan. Savršeno precizan odgovor na pogrešno postavljen zadatak je sigurno lošiji od procene dobro postavljenog zadatka.

Prilikom dostavljanja prognoziranih podataka krajnjem korisniku dodaje se i preporuka da se vrednosti uzimaju sa određenom rezervom. Prognoze kao proizvod sadašnjosti i prošlosti i sklone su promenama na koje se ne može uticati, sa jedne, i sa druge strane sa dozom subjektivnosti i slobode autora. Autor ili ekspert ima slobodu da pojavljivanje "nepredvidivih" događaja tretira kako želi, odakle se može zaključiti da uvek postoji dovoljno dobra prognoza.

"Ako mislite da statistika nema ništa da kaže o onome što radite ili kako da to radite bolje, onda ili grešite ili je vreme da nađete interesantniji posao": Stephen Senn[4].

[4] Stephen Senn (1963-) engleski statističar

7 LITERATURA

[1] Pavle Mladenović, *Verovatnoća i statistika,* Matematički fakultet, Beograd 2008.

[2] Christian Francq, Jean-Michel Zakoan, *GARCH Models : Structure, Statistical Inference and Financial Applications,* John Wiley & Sons Ltd 2010.

[3] Zlatko J. Kovačić, *Analiza vremenskih serija*, Ekonomski fakultet, Beograd 1995.

[4] Alexander J. McNeil, Rüdiger Frey, Paul Embrechts, *Quantitative Risk Management:Concepts, Techniques and Tools,* by Princeton University Press 2005.

[5] George E.P. Box, Gwilym M. Jenkins, Gregory C. Reinsel, *Time Series Analysis – Forecasting and Control (Fourt Edition)*, Statistics Department, United States od America 2008.

[6] Z. Mladenović, A. Nojković, *Primenjena analiza vremenskih serija*, Ekonomski fakultet, Beograd 2012.

[7] William W.S. Wei, Times Series Analysis – *Univariate and Multivariate Methods (Second Edition)*, Department of Statistics, The Fox School of Business and Management, Temple University, 2006.

[8] Jovan Mališić, Vesna Jevremović, *Slučajni procesi i vremenske serije*, Beograd 2008.

[10] Nebojša Marić, *Modeliranje vremenskih serija*, Savezni zavod za statistiku, Beograd 1991.

[11] Baza vremenskih serija Republičkog zavoda za statistiku Srbije pod SQL serverom i IST aplikacijom.

[12] Nada Đerić, *Upustvo za korištenje softvera JDemetra+*, Republički zavod za statistiku, Beograd 2014.

[13] Spyros Makridakis, Steven C. Wheelwright, *Interactive Forecasting*, by Holden-Day, Inc., San Francisco 1978.