

UNIVERZITET U BEOGRADU  
МАТЕМАТИЧКИ ФАКУЛТЕТ У БЕОГРАДУ



# ДЕСЕЗОНИРАЊЕ ВРЕМЕНСКИХ СЕРИЈА АЛАТОМ JDemetra+

— MASTER RAD —

**Mentor:**

PROF. DR VESNA JEVREMOVIĆ

**Kandidat:**

MONIKA MLADENOVIĆ

BEOGRAD, JUN 2015. GODINE

U STVARIMA KOJE SE TIČU SAZNANJA,  
PRISUSTVO *dobra* NE DAJE JEDINO TO DA  
BUDU SAZNATE, NEGO UPRAVO IZ NJEGA  
SAMOG PROISHODI NJIHOVO  
BIVSTVOVANJE I NJIHOVA SUŠTINA, A  
SAMO *dobro* NIJE SUŠTINA, NEGO SE PO  
UZVIŠENOSTI I MOĆI UZDIŽE IZNAD NJE.

Platon: „Država”

## REZIME

Vremenska serija, definisana kao niz opservacija vezanih za jednu pojavu i uređenih u odnosu na vreme, za osnovni cilj ima da na osnovu istorijskih podataka prognozira buduću vrednost te pojave. Za nalaženje objašnjenja o uzročno-posledičnoj vezi opservacija i načina za predviđanje kretanja serije u narednom periodu, koriste se različiti modeli analize, a sve u cilju kontrolisanja posmatrane pojave manipulacijom veličina koje na nju utiču. S obzirom na to da su vremenske serije podležne uticaju sezone, za pravilno razumevanje prirode serije i njenog kretanja kroz vreme neophodan je proces desezoniranja.

Jedan od načina za uklanjanje sezonske komponente je metod *X-13-ARIMA*, koji spada u kategoriju metoda empirijskog filtriranja, a na kome je zasnovano funkcionisanje *open source* alata *JDemetra+* (pored *TramoSeats* metode). Osnova ove metode je računarski program *X-11*, najpoznatija operacionalizacija dekompozicije, koji se koristi i od strane zvanične statistike. *JDemetra+ (ver.2.0.0)*, koja će u radu biti korišćena, javno je dostupna od januara 2015. godine, a zbog svog kvaliteta i široke primene u analizi vremenskih serija, koristi se i u vodećim evropskim zemljama.

Osim metodološkog aspekta, ovaj rad ima za zadatak da pokaže praktičnu implementaciju teorijskih pojmove i konkretno, na određenim serijama Republike Srbije, prikaže na koji način se bira specifikacija za najkvalitetnije raščlanjivanje originalne serije na komponente (od uvođenja nacionalnog kalendarja, preko transformacija podataka, do odabira modela, detektovanja autlajera i drugo). Spektar mogućnosti koje ovaj alat nudi je ogroman, te je osnovni cilj rada stvoriti uvid u reprezentaciju i analizu 'živih' podataka i dati praktičan osvrt na ovu široko primenjenu oblast analize vremenskih serija.

## ABSTRACT

Time series, defined as a sequence of observations which are related to one occurrence and ordered in time, for its basic aim has the forecasting of future value of that phenomenon based on historical data. To find explanation of the causal connection between the observations and method for predicting the movement of the series in the future period of time, we use different models of analysis, with the aim of controlling observed phenomenon, manipulating variables that affect it. Due to the fact that the time series are susceptible to the influence of the season, for a proper understanding of the nature of the series and its movement through time, seasonal adjustment process is necessary.

One of the ways for elimination of the seasonal component is a method *X-13-ARIMA*, which belongs to the category of empirical filtering methods, and on which is based the functioning of open source tool *JDemetra+* (besides *TramoSeats* method). The basis of this method is a computer program *X-11*, best known operationalization of decomposition, which is used by official statistics. *JDemetra+ (ver.2.0.0)*, which will be used in this thesis, is publicly available since January 2015, and because of its quality and widely utilization in time series analysis, it is used in leading European countries.

Apart from the methodological point of view, this thesis has the task to show the practical implementation of theoretical concepts and specifically, on certain series of the Republic of Serbia, show how to choose specification for best parsing of original series on components (from introduction of the national calendar, through the transformation of data, to the selection of model, detection of outliers, etc.). The spectrum of possibilities that this tool offers is huge, and the main aim of this thesis is to provide insight into the representation and analysis of 'live' data and to give a practical overview of this widely applied field of time series analysis.

# SADRŽAJ

<b>1 STOHAŠTIČKA PRIRODA VREMENSKIH SERIJA</b>	<b>1</b>
1.1 STOHAŠTIČKI PROCESI - POJMOVI I DEFINICIJE . . . . .	1
1.2 AUTOKOVARIJACIONA, AUTOKORELACIONA I PARCIJALNA AUTOKORELACIONA FUNKCIJA STACIONARNOG STOHAŠTIČKOG PROCESA . . . . .	3
1.3 DEFINICIJA I KLASIFIKACIJA VREMENSKIH SERIJA . . . . .	5
1.4 ISTORIJSKI RAZVOJ ANALIZE VREMENSKIH SERIJA . . . . .	7
1.5 CILJEVI ANALIZE VREMENSKIH SERIJA . . . . .	9
1.6 OPERATORI U ANALIZI VREMENSKIH SERIJA . . . . .	10
<b>2 STACIONARNE VREMENSKE SERIJE</b>	<b>11</b>
2.1 BELI ŠUM . . . . .	11
2.2 LINEARNI PROCESI . . . . .	12
2.3 AUTOREGRESIONI PROCESI REDA $p$ ( $AR(p)$ PROCESI) . . . . .	14
2.3.1 AUTOREGRESIONI PROCESI PRVOG REDA ( $AR(1)$ ) . . . . .	16
2.3.2 AUTOREGRESIONI PROCESI DRUGOG REDA ( $AR(2)$ ) . . . . .	17
2.4 PROCESI POKRETNIH PROSEKA REDA $q$ ( $MA(q)$ PROCESI) . . . . .	19
2.4.1 PROCESI POKRETNIH PROSEKA PRVOG REDA ( $MA(1)$ ) . . . . .	20
2.4.2 PROCESI POKRETNIH PROSEKA DRUGOG REDA ( $MA(2)$ ) . . . . .	21
2.5 AUTOREGRESIONI PROCESI POKRETNIH PROSEKA ( $ARMA(p, q)$ PROCESI)	22
<b>3 NESTACIONARNE VREMENSKE SERIJE</b>	<b>24</b>
3.1 STABILIZACIJA NESTACIONARNOSTI U ODNOSU NA SREDNU VREDNOST . . . . .	24
3.2 STABILIZACIJA NESTACIONARNOSTI U ODNOSU NA VARIJANSU . . . . .	26
3.3 $ARIMA(p, d, q)$ PROCESI . . . . .	27
3.4 IZGRADNJA $ARIMA$ MODELA . . . . .	28
3.5 SEZONSKI $ARIMA$ MODELI . . . . .	29
3.5.1 $ARIMA(0, 1, 1)(0, 1, 1)$ MODEL . . . . .	31
<b>4 DESEZONIRANJE VREMENSKIH SERIJA</b>	<b>32</b>
4.1 UVOD . . . . .	32
4.2 PREDNOSTI, MERE OPREZA, TROŠKOVI I RIZICI DESEZONIRANJA . . . . .	33
4.3 ISTORIJA PROCESA DESEZONIRANJA . . . . .	34
4.4 POSTUPAK DESEZONIRANJA . . . . .	37
4.4.1 CILJEVI PRED-PODEŠAVANJA SERIJE (ENGL. <i>Pre-adjustment</i> ) . . . . .	37
4.4.2 GRAFIČKA ANALIZA SERIJE . . . . .	38
4.4.3 KALENDARSKO PRILAGOĐAVANJE . . . . .	39
4.4.4 PRONALAŽENJE I ISPRAVLJANJE EKSTREMNIH VREDNOSTI (ENGL. <i>Outliers</i> ) . . . . .	41
4.4.5 SHEMA DEKOMPOZICIJE . . . . .	42
4.4.6 ODABIR MODELA . . . . .	46
4.4.7 DIREKTNI/INDIREKTNI PRISTUP DESEZONIRANJU . . . . .	52
4.4.8 SPECIFIČNI PROBLEMI DESEZONIRANJA . . . . .	53
4.4.9 PROGNOZA (ENGL. <i>Forecasting</i> ) . . . . .	54
4.4.10 REVIZIJA . . . . .	54
4.4.11 KVALITET DESEZONIRANJA . . . . .	56

<b>5 ALAT ZA DESEZONIRANJE - JDemetra+ (2.0.0)</b>	<b>58</b>
5.1 PREGLED JDEMETRE+ . . . . .	58
5.2 IMPORTOVANJE SERIJE/GRUPE SERIJA . . . . .	59
5.3 KREIRANJE I IMPORTOVANJE NACIONALNOG KALENDARA . . . . .	60
5.4 KREIRANJE SPECIFIKACIJE ZA PROCES DESEZONIRANJA . . . . .	65
5.4.1 SKRAĆIVANJE SERIJE (ENGL. <i>Model/Series span</i> ) . . . . .	67
5.4.2 TRANSFORMACIJA SERIJE . . . . .	68
5.4.3 KALENDAR . . . . .	68
5.4.4 AUTLAJERI . . . . .	68
5.4.5 ARIMA MODEL . . . . .	69
5.4.6 SEZONSKI FILTER . . . . .	70
5.5 OBRADA JEDNE SERIJE (ENGL. <i>Single processing</i> ) . . . . .	71
5.6 OBRADA GRUPE SERIJA (ENGL. <i>Multi processing</i> ) . . . . .	79
5.7 KREIRANJE IZLAZNOG FAJLA SA REZULTATIMA (ENGL. <i>Output</i> ) . . . . .	81
5.8 REVIZIJA PROCESA DESEZONIRANJA (ENGL. <i>Refresh options</i> ) . . . . .	83
<b>6 PRIMERI KORIŠĆENJA ALATA JDemetra+ (2.0.0)</b>	<b>85</b>
6.1 POREĐENJE UTICAJA KREIRANE I <i>default</i> SPECIFIKACIJE . . . . .	85
6.2 UTICAJ AUTLAJERA NA LINIJU TRENDI . . . . .	91
6.3 POREĐENJE UTICAJA NACIONALNOG I <i>default</i> KALENDARA . . . . .	93
6.4 POREĐENJE X13 i TramoSeats METODA KOD NESEZONSKE SERIJE . . . . .	95
<b>7 ZAKLJUČAK</b>	<b>98</b>
<b>8 LITERATURA</b>	<b>99</b>

# 1 STOHALSTIČKA PRIRODA VREMENSKIH SERIJA

Vremenske serije su odraz prirode ljudske svakodnevnice, kao posledica uređenosti uslovljene vremenom. U ekonomiji se, na primer, prate godišnje vrednosti društvenog proizvoda, kvartalni podaci o raspoloživim sredstvima domaćinstava, mesečna kretanja zaliha gotovih proizvoda u industrijskoj proizvodnji ili dnevne fluktuacije deviznog kursa. U demografiji se govori o godišnjim stopama nataliteta i mortaliteta, u neurofiziologiji mere se električne aktivnosti mozga u toku psihičkih procesa (takva vremenska serija odražava amplitudu tog signala u vremenu), a u geofizici prate se učestalosti aktivnosti zemlje. Poljoprivrednike zanima godišnji prinos proizvoda i kretanje otkupnih i prodajnih cena, a meteorologe prosečne mesečne ili godišnje padavine na određenoj teritoriji.

Formiranje vremenske serije na osnovu zabeleženih numeričkih podataka u toku izvesnog perioda posmatranja omogućava svaku dalju matematičku analizu. U cilju definisanja teorijskih osnova savremenih statističkih modela vremenskih serija, kao i okvira analize istih, najpre treba uvesti pojam stohastičkog procesa.

## 1.1 Stohastički procesi - pojmovi i definicije

**Definicija 1.1.1:** [11]  $\Omega$  je skup ishoda jednog eksperimenta. Elementi ovog skupa označavaju se sa  $\omega_i$ ,  $i = 1, 2, \dots$  i nazivaju *elementarni događaji*.

**Definicija 1.1.2:** [11] Podskup  $\mathcal{F}$  partitivnog skupa  $\mathcal{P}(\Omega)$  je  $\sigma$ -polje ( $\sigma$ -algebra) nad  $\Omega$ , ako važe uslovi:

- $\Omega \in \mathcal{F}$  (1.1.1a)

- $A \in \mathcal{F} \Rightarrow \bar{A} \in \mathcal{F}$  (1.1.1b)

- $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$  (1.1.1c)

**Definicija 1.1.3:** [11] Neka je  $\Omega$  skup svih elementarnih događaja i  $\mathcal{F}$   $\sigma$ -polje nad  $\Omega$ . Funkcija  $P : \mathcal{F} \rightarrow [0, 1]$  zove se *verovatnoća* na prostoru  $(\Omega, \mathcal{F})$  ako zadovoljava uslove:

- $P(\Omega) = 1$  (1.1.2a)

- $P(A) > 0, \forall A \in \mathcal{F}$  (1.1.2b)

- $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{F}, A_i \cap A_j = \emptyset, i \neq j, i = 1, 2, \dots \Rightarrow P\left(\sum_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$  (1.1.2c)

**Definicija 1.1.4:** [11] Uređena trojka  $(\Omega, \mathcal{F}, P)$  zove se *prostor verovatnoća*.

**Definicija 1.1.5:** [11] Neka je  $(\Omega, \mathcal{F}, P)$  prostor verovatnoća. Preslikavanje  $X : \Omega \rightarrow \mathbb{R}$  zove se *slučajna promenljiva* ako za  $\forall B \in \mathcal{B}$  važi da je  $X^{-1}(B) \in \mathcal{F}$ , gde je  $\mathcal{B}$  Borelova  $\sigma$ -algebra. Za  $X$  kažemo da je  $\mathcal{F}$ -merljivo.

Do pojma stohastičkog procesa dolazimo proširujući pojam slučajne promenljive. Slučajna promenljiva je funkcija mogućih ishoda statističkog eksperimenta. Ako posmatramo njenu evoluciju tokom vremena, ona postaje i funkcija vremena. U opštem slučaju ne zavisi od

vremena, međutim, mnoge pojave i događaji, čiji su ishodi neizvesni, a koje se odvijaju u vremenu, zahtevaju da se pojam slučajne promenljive uopšti tako da se uključi i vremenska komponenta ( $t$ ). Posmatrajući familiju slučajnih promenljivih koje zavise od vremena, dolazimo do definicije *stohastičkog ili slučajnog procesa*.

**Definicija 1.1.6:** [12] Familija slučajnih promenljivih  $\{X_t : t \in T, \omega \in \Omega\}$  definisanih nad istim prostorom verovatnoće  $(\Omega, \mathcal{F}, P)$  zove se *stohastički proces* sa indeksnim skupom  $T$ .

Promenljiva  $\omega$  se često izostavlja u zapisu, pa se umesto toga stohastički proces označava sa  $\{X_t : t \in T\}$ . Očigledno je da proces zavisi od dve promenljive -  $t$  i  $\omega$ . Za fiksirano  $t_0 \in T$ , stohastički proces je jedna slučajna promenljiva. Za fiksirano  $\omega_0 \in \Omega$ ,  $X_t(\cdot)$  je funkcija definisana na skupu  $T$ , koja se naziva realizacija ili *trajektorija* stohastičkog procesa. Za fiksirane  $t_0 \in T$  i  $\omega_0 \in \Omega$ ,  $X_t$  je realan broj ili jedna realizacija stohastičkog procesa.

**Definicija 1.1.7:** [12] Neka je  $\mathcal{T}$  skup  $((t_1, t_2, \dots, t_n) \in T^n, t_1 < t_2 < \dots < t_n, n = 1, 2, \dots)$ . Tada su *konačno dimenzionalne funkcije raspodele* stohastičkog procesa funkcije  $\{F_t(\cdot), t \in \mathcal{T}\}$  definisane sa

$$F_t(x) = P\{X_{t_1} \leq x_1, X_{t_2} \leq x_2, \dots, X_{t_n} \leq x_n\}, \quad (x_1, x_2, \dots, x_n) \in \mathbb{R}^n. \quad (1.1.3)$$

Sistem funkcija raspodele  $\{F_t(\cdot), t \in \mathcal{T}\}$  zadovoljava dva uslova:

- *uslov simetrije* - ako je  $\{i_1, \dots, i_n\}$  jedna permutacija brojeva 1 do  $n$ , tada važi

$$F_{t_{i_1}, \dots, t_{i_n}}(x_{i_1}, \dots, x_{i_n}) = F_{t_1, \dots, t_n}(x_1, \dots, x_n). \quad (1.1.4a)$$

- *uslov kompatibilnosti* - ako je  $m < n$  za proizvoljne  $t_{m+1}, \dots, t_n \in T^n$ , tada važi

$$F_{t_1, \dots, t_m, t_{m+1}, \dots, t_n}(x_1, \dots, x_m, \infty, \dots, \infty) = F_{t_1, \dots, t_m}(x_1, \dots, x_m). \quad (1.1.4b)$$

**TEOREMA 1.1.1:** (Kolmogorov) [12] Za svaku familiju funkcija raspodele koje zadovoljavaju uslove simetrije i kompatibilnosti postoji prostor verovatnoće  $(\Omega, \mathcal{F}, P)$  i stohastički proces  $\{X_t : t \in T\}$  definisan na njemu, koji ima date raspodele kao svoje konačno-dimenzionalne raspodele.

Neka je  $\{X_t : t \in T\}$  stohastički proces. Tada je:

- srednja vrednost stohastičkog procesa:

$$E(X_t) = \mu_t, \quad t \in T, \quad (1.1.5a)$$

- varijansa stohastičkog procesa:

$$\text{Var}(X_t) = \sigma_t^2, \quad t \in T, \quad (1.1.5b)$$

- kovarijansa stohastičkog procesa:

$$\text{Cov}(X_r, X_s) = E((X_r - E(X_r))(X_s - E(X_s))), \quad r, s \in T, \quad (1.1.5c)$$

- korelacija stohastičkog procesa:

$$\rho_x(r, s) = \frac{\text{Cov}(X_r, X_s)}{\sqrt{\text{Var}(X_r)\text{Var}(X_s)}}, \quad r, s \in T. \quad (1.1.5d)$$

**Definicija 1.1.8:** [12] Stohastički proces je *stogo stacionaran* ako su njegove konačno dimenzionalne raspodele invarijantne u odnosu na  $t$ , odnosno ako za  $t_i, t_i + t \in T$ ,  $i = 1, 2, \dots$  važi:

$$F_{t_1, \dots, t_n}(x_1, \dots, x_n) = F_{t_1+t, \dots, t_n+t}(x_1, \dots, x_n). \quad (1.1.6)$$

**Definicija 1.1.9:** [12] Stohastički proces  $\{X_t, t \in T\}$  je *slabo stacionaran* ako su ispunjeni sledeći uslovi:

- $E(X_t) = \mu = \text{const}, \forall t \in T$  (1.1.7a)

- $E(X_t^2) < \infty, \forall t \in T$  (1.1.7b)

- $\text{Cov}(X_r, X_s) = \text{Cov}(X_{r+t}, X_{s+t}) = E[(X_{r+t} - \mu)(X_{s+t} - \mu)], \forall r, s, t \in T$  (1.1.7c)

Treći uslov podrazumeva da je kovarijansa stacionarnog slučajnog procesa samo funkcija udaljenosti promenljivih ( $\gamma_{r,s} = \gamma_{r+t,s+t} = \gamma_{r-s,0} = \gamma_\tau, \forall r, s, t, r+t, s+t, \tau \in T$ ), odnosno ona ne zavisi od perioda u kojem se razmatraju izabrane promenljive, već samo od njihovog međusobnog rastojanja u vremenu. *Iz stroge stacionarnosti uvek sledi slaba stacionarnost.*

**Definicija 1.1.10:** [12] Stohastički proces je *Gausovski* ako sve njegove konačno dimenzionalne funkcije raspodele imaju višedimenzionalnu normalnu raspodelu.

Ako je  $\{X_t, t \in T\}$  slabo stacionaran Gausovski proces, onda slučajni vektori  $(X_{t_1}, \dots, X_{t_n})$  i  $(X_{t_1+h}, \dots, X_{t_n+h})$ ,  $\forall h, t_1, t_2, \dots \in \mathbb{Z}$  i  $\forall n \in \mathbb{N}$  imaju istu očekivanu vrednost i autokovarijansnu matricu, pa je onda  $\{X_t, t \in T\}$  stogo stacionaran proces, što dalje implicira da jedino u slučaju kada je stohastički proces Gausovski, slaba stacionarnost implicira strogu, a obrnuto ne važi.

Dakle, stacionarnost stohastičkog procesa podrazumeva da se raspodela verovatnoća proizvoljno odabranog slučajnog vektora sa  $n \in \mathbb{N}$  komponenata ne menja duž vremenske ose. S obzirom na to da je taj uslov prestrog, a i teško ga je ispitati, on se zamenuje slabijim uslovom o stacionarnosti drugog reda, tj. uslovom slabe stacionarnosti. Nadalje će se koristiti termin *stacionaran* kada se misli na slabu stacionarnost, osim kada se naglasi drugačije.

## 1.2 Autokovarijaciona, autokorelaciona i parcijalna autokorelaciona funkcija stacionarnog stohastičkog procesa

Prema definiciji 1.1.9, kod stacionarnog procesa  $X_t$  sredina i varijansa su konstantne, a kovarijanse su funkcije samo od vremenske razlike. Zato je kovarijansa između  $X_t$  i  $X_{t+k}$  data sledećom definicijom:

**Definicija 1.2.1:** [12] Funkcija stacionarnog stohastičkog procesa  $\{X_t, t \in T\}$  definisana sa

$$\gamma_k = \text{Cov}(X_t, X_{t+k}) = E((X_t - E(X_t))(X_{t+k} - E(X_{t+k}))), \quad t \in T, k \in \mathbb{Z} \quad (1.2.1)$$

naziva se *autokovarijaciona funkcija*.

**Definicija 1.2.2:** [12] Koeficijent korelacije između  $X_t$  i  $X_{t+k}$  je:

$$\rho_k = \frac{\text{Cov}(X_t, X_{t+k})}{\sqrt{\text{Var}(X_t)} \sqrt{\text{Var}(X_{t+k})}} = \frac{\gamma_k}{\gamma_0}, \quad t \in T, k \in \mathbb{Z}, \quad (1.2.2)$$

gde je  $\text{Var}(X_t) = \text{Var}(X_{t+k}) = \gamma_0$ . Funkcija  $\rho_k$  naziva se *autokorelaciona funkcija*.

**Definicija 1.2.3:** [12] Kovarijaciona i korelaciona matrica reda  $n$  definisane su, redom, na sledeći način:

$$\Gamma_n = \begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{n-1} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{n-1} & \gamma_{n-2} & \cdots & \gamma_0 \end{bmatrix} \quad R_n = \begin{bmatrix} 1 & \rho_1 & \cdots & \rho_{n-1} \\ \rho_1 & 1 & \cdots & \rho_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n-1} & \rho_{n-2} & \cdots & 1 \end{bmatrix} \quad (1.2.3)$$

Očigledno je da važi  $\Gamma_n = \sigma^2 R_n$ .

**TEOREMA 1.2.1:** [17] Neka je  $\gamma_k$  i  $\rho_k$ ,  $k \in \mathbb{Z}$  autokovarijaciona i autokorelaciona funkcija slabo stacionarnog stohastičkog procesa. Tada važi:

$$\bullet \gamma_0 = \text{Var}(X_t) \geq 0, \quad \rho_0 = 1 \quad (1.2.4a)$$

$$\bullet \gamma_k = \gamma_{-k}, \quad \rho_k = \rho_{-k} \quad (1.2.4b)$$

$$\bullet |\gamma_k| \leq \gamma_0, \quad |\rho_k| \leq 1 \quad (1.2.4c)$$

U statistici se često dešava da je korelacija između dve promenljive u stvari rezultat njihove korelisanosti sa trećom promenljivom u modelu. Zbog toga se uvodi pojam *parcijalne korelacije*. Parcijalna korelacija predstavlja korelaciju između dve promenljive uz eliminisan uticaj drugih promenljivih iz modela.

**Definicija 1.2.4:** [12] *Parcijalni koeficijent korelacije* između  $X_t$  i  $X_{t-k}$ , u oznaci  $\phi_{kk}$ , je  $k$ -ti regresioni koeficijent u autoregresionom procesu reda  $k$ . Niz  $\phi_{11}, \phi_{22}, \dots$  predstavlja parcijalnu autokorelacionu funkciju, čiji se grafički prikaz naziva *parcijalni koreogram*. Parcijalni autokorelacioni koeficijent  $\phi_{kk}$  definiše se kao obični autokorelacioni koeficijent između formiranih članova  $(X_t - \hat{X}_t)$  i  $(X_{t-k} - \hat{X}_{t-k})$ :

$$\phi_{kk} = \frac{\text{Cov}[(X_t - \hat{X}_t), (X_{t-k} - \hat{X}_{t-k})]}{\sqrt{\text{Var}(X_t - \hat{X}_t)\text{Var}(X_{t-k} - \hat{X}_{t-k})}}, \quad (1.2.5)$$

gde je  $\hat{X}_t$  ocena zavisnosti  $X_t$  od  $X_{t-1}, X_{t-2}, \dots, X_{t-k+1}$ , dobijena metodom najmanjih kvadrata, koja opisuje kretanje  $X_t$  u funkciji datog skupa objašnjavajućih promenljivih. Dakle, razlika  $(X_t - \hat{X}_t)$  predstavlja deo od  $X_t$  koji ne sadrži uticaj  $X_{t-1}, X_{t-2}, \dots, X_{t-k+1}$ . Analogno važi i za  $(X_{t-k} - \hat{X}_{t-k})$ .

**Definicija 1.2.5:** [17] Funkcija  $k : \mathbb{Z} \rightarrow \mathbb{R}$  je *nenegativno definitna* ako

$$\sum_{i,j=1}^n a_i k(t_i - t_j) a_j \geq 0, \quad \forall n \in \mathbb{N}, \quad \forall a = (a_1, \dots, a_n) \in \mathbb{R}^n \text{ i } \forall t = (t_1, \dots, t_n). \quad (1.2.6)$$

**TEOREMA 1.2.2:** [17] Ako je  $\{X_t, t \in T\}$  slabo stacionaran proces, onda je njegova kovarijaciona funkcija nenegativno definitna.

### 1.3 Definicija i klasifikacija vremenskih serija

U literaturi ne postoji jedinstven stav oko toga šta je vremenska serija. Sledeća dva mišljenja su dominantna [10]:

1. Vremenska serija predstavlja jednu realizaciju stohastičkog procesa. U tom smislu odnos vremenske serije i slučajnog procesa odgovara odnosu uzorka i osnovnog skupa u standardnoj teoriji statističkog zaključivanja. Kao što uzorak predstavlja deo osnovnog skupa na osnovu koga se izvode zaključci o svojstvima osnovnog skupa, tako i analiza konkretnе vremenske serije treba da omogući sagledavanje karakteristika slučajnog procesa.
2. Ne postoji razlika između slučajnog procesa i vremenske serije. To znači da možemo smatrati da vremenska serija predstavlja uređeni niz opservacija. Drugim rečima, termin slučajnog procesa i vremenske serije su sinonimi. Uređivanje se najčešće, ali ne nužno, vrši u odnosu na vreme<sup>1</sup> i to obično u jednakim vremenskim intervalima.

Na osnovu svega prethodnog, u daljem tekstu naredna definicija smatraće se relevantnom:

**Definicija 1.3.1:** [7] *Vremenska serija* je hronološki uređen niz vrednosti neke pojave, tj. posmatrani niz vrednosti procesa koji generiše pojavu u vremenu, pri čemu se vrednosti odnose na jednakе vremenske intervale, tj. vezane su za ekvidistantne tačke na vremenskoj osi.

Prema prirodi posmatrane pojave, razlikujemo dva tipa vremenskih serija:

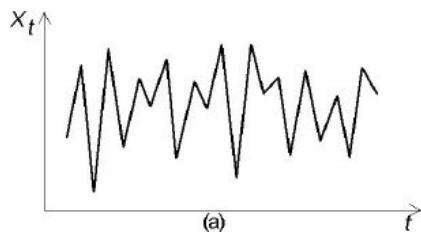
- *neprekidne* - one kod kojih se opservacije mogu registrovati u ma kom vremenskom intervalu;
- *prekidne* - one kod kojih se opservacije mogu beležiti u istim vremenskim intervalima (dnevno, mesečno, kvartalno ili godišnje).

Prekidna vremenska serija može se dobiti od neprekidne, tako što se kod neprekidne snimanje posmatrane pojave izvrši samo u određenim (jednakim) vremenskim intervalima (*sistematskim uzorkom*) ili posmatranjem neprekidne vremenske serije u izabranom intervalu vremena (*vremenskim agregiranjem*).

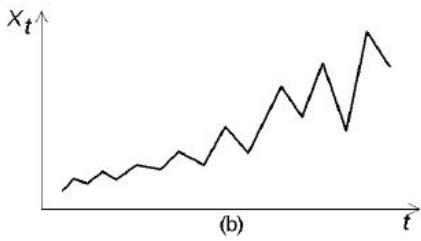
Neka je sa  $X_t$  označena opservacija vremenske serije u trenutku  $t$ . Na slikama 1.1 i 1.2 predstavljeni su najčešći oblici vremenskih serija:

Slika 1.1: Primeri najčešćih oblika vremenskih serija, [7]

- (a) *stacionarna u odnosu na srednju vrednost*  
- nivo serije ne menja se tokom vremena,  
tj. karakteristična su odstupanja oko konstantnog nivoa;

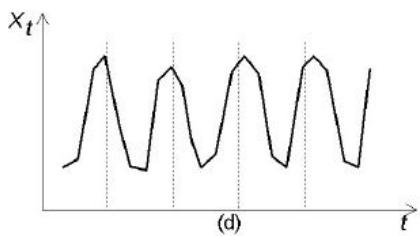


1. Postoje i *prostorne* ili *geografske* serije kod kojih se uređivanje vrši u odnosu na prostor.

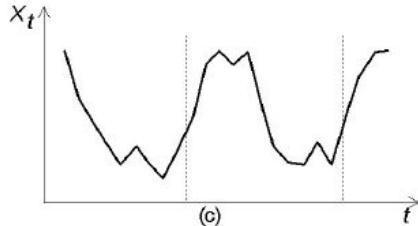


(b) serija sa uzlaznim *trendom* i slučajnim fluktuacijama oko njega - nestacionarna serija u odnosu na srednju vrednost (nivo serije se povećava tokom vremena), a kako sa porastom nivoa raste i varijansa, nestacionarna je i u odnosu na varijansu;

- (c) serija sa *cikličnim varijacijama* - period ponavljanja sličnog toka serije duži je od godinu dana;

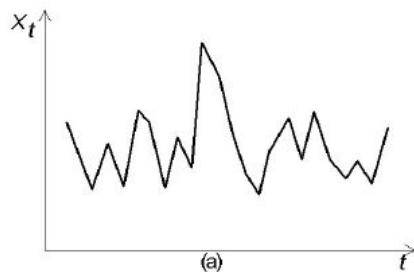
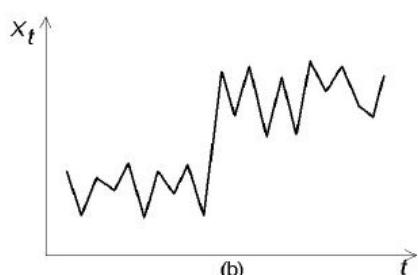


(d) serija sa *sezonskim varijacijama* - oscilacije se ponavljaju u periodima do godinu dana (mesečno, kvartalno ili polugodišnje).



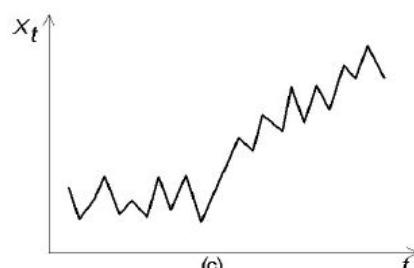
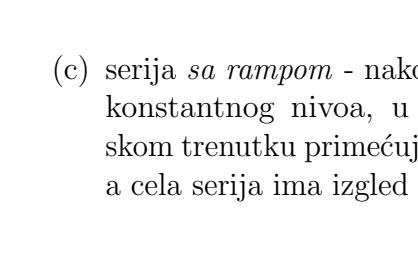
**Slika 1.2:** Tipovi vremenskih serija čija se struktura tokom vremena jednokratno ili trajno promenila pod dejstvom egzogenih poremećaja, [7]

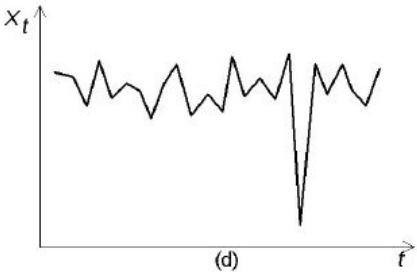
- (a) serija sa *jednokratnom intervencijom* - kad u jednom trenutku dođe do promene u strukturi serije, usled spoljnog, jednokratnog poremećaja, nekoliko (na slici tačno dve) opservacija promeni nivo u odnosu na ostale;



- (b) serija sa *stepenastom intervencijom* - kad u izvesnom vremenskom trenutku dođe do trajnog porasta nivoa serije, to se odražava na stepenast izgled njenog nivoa;

- (c) serija sa *rampom* - nakon fluktuacija oko konstantnog nivoa, u jednom vremenskom trenutku primećuje se uzlazni trend, a cela serija ima izgled podignute rampe;

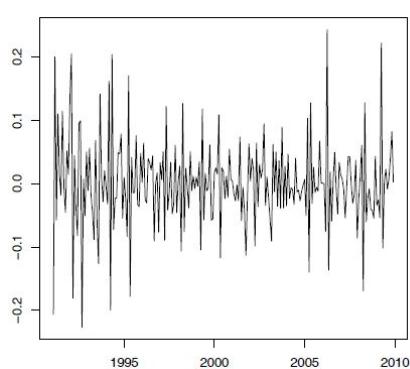




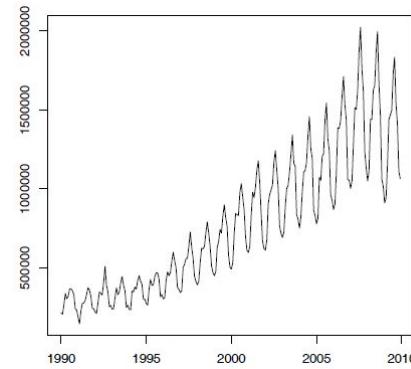
- (d) serija sa *nestandardnom opservacijom* - sadrži jednu nestandardnu opservaciju (engl. *outlier*), koja je posledica nekog ekstremnog uslova ili je čak slučajna greška u registrovanju posmatrane pojave.

Za razliku od klasične statističke analize gde su elementi slučajnog uzorka međusobno nezavisni, kod analize vremenskih serija opservacije u uzorku su međusobno zavisne, samim tim što se u obzir uzima njihov vremenski poredak. Upravo se ta zavisnost koristi za formiranje modela vremenske serije, u cilju prognoziranja budućih opservacija na osnovu prethodnih.

Klasična ekonometrijska analiza vremenskih serija zasniva se na pretpostavci da raspoloživi podaci poseduju svojstvo stacionarnosti, pod kojim se, slobodno rečeno, podrazumeva vremenska serija čije se kretanje tokom vremena odvija po ustaljenom obrascu u smislu ne-promenljivosti njegovih svojstava, tj. podrazumeva konstantnost srednje vrednosti i varijanse vremenske serije tokom vremena. Najveći broj ekonomskih vremenskih serija ne zadovoljava uslove stacionarnosti, tj. menjaju se u toku vremena, što dovodi do kretanja mnogih kvalitativnih i kvantitativnih veličina. Te promene se najčešće ispoljavaju kroz nestabilnost srednje vrednosti i varijanse [17].



(a) *Stacionarna serija*



(b) *Nestacionarna serija*

**Slika 1.3:** Tipovi vremenskih serija u odnosu na stacionarnost

Analiza vremenskih serija podrazumeva upoznavanje tog stohastičkog procesa modeliranjem relevantnih promenljivih veličina i njihovih odnosa. Konvergenciju ekonometrijskog i statističkog pristupa analizi, nastalih u poslednjoj deceniji, najlakše je razumeti osvrtanjem na istorijski razvoj analize vremenskih serija.

## 1.4 Istorijski razvoj analize vremenskih serija

Vremenske serije su imale značajnu ulogu još u davnoj fazi razvoja prirodnih nauka. Prvobitno je analiza vremenskih serija bila orijentisana na analizu onih serija koje su se javljale u astronomiji, geofizici i meteorologiji. Poznato je da su još vavilonski astronomi koristili vremenske serije relativnih položaja zvezda i planeta da predvide astronomiske događaje. Analiza vremenskih serija odvajkada doprinosi otkrivanju pravilnosti u opservacijama promenljivih i izvođenju 'zakona' iz njih, kao i u predviđanju budućih događaja (npr. uočavano je relativno glatko kretanje vremenskih serija sa cikličnim pravilnostima, kao što

su jedanaestogodišnji ciklusi u slučaju sunčanih pega ili godišnji ciklusi u meteorološkim podacima). Sa metodološke strane gledišta, glavni pravac istraživanja bio je usmeren na otkrivanje 'skrivenih perioda' u vremenskoj seriji. Osnovna ideja, poznata još kod astronoma sedamnaestog veka, bila je mogućnost raščlanjivanja vremenske serije na konačan broj nezavisnih, ali ne direktno registrovanih komponenti, koje se redovno dobijaju, i kao takve mogu unapred biti izračunate. Opservacije kretanja planeta formirale su osnovu i za formulisanje Keplerovih zakona (*Johannes Kepler*).

Krajem devetnaestog i početkom dvadesetog veka glavni doprinos u pogledu analize skrivenih perioda daje *Sir Arthur Schuster* (1898) definisanjem tzv. periodograma. U savremenoj analizi prvobitna namena periodograma je uglavnom napuštena, ali se koristi kao polazno sredstvo za ocenu spektralne funkcije gustine. Sredinom devetnaestog veka, ovaj metodološki pristup astronomiji preuzeli su ekonomisti *Charles Babbage* i *William Stanley Jevons*. Raščlanjivanje na komponente, koje zavisi od različitih uzročnih faktora, što je obično i slučaj u klasičnim analizama vremenskih serija, uveo je *Warren M. Persons* (1919). On ističe četiri različite komponente:

- dugoročni razvoj - trend,
- cikličnu komponentu sa periodom većim od godinu dana - ciklus,
- komponentu koja sadrži uspone i padove tokom godine - sezona i
- komponentu koja sadrži sva kretanja koja nisu deo ni trenda, niti ciklusa ili sezone - rezidual.

Pod pretpostavkom da su različiti faktori nezavisni, njihovo aditivno preslojavanje generiše vremensku seriju, koju jedino kao celinu i možemo da opažamo. Da bi se dobole informacije o procesu generisanja podataka, potrebno je napraviti određene pretpostavke o komponentama. Za klasičnu vremensku seriju pretpostavlja se da na njene sistematske komponente (trend, ciklus i sezonu) ne utiču stohastički poremećaji, te stoga mogu biti predstavljeni determinističkim funkcijama vremena. Stohastički uticaj je ograničen na reziduale, koji ne sadrže nikakva sistematska kretanja. Oni su stoga modelirani kao serije nezavisnih ili nekorelisanih slučajnih veličina sa očekivanjem nula i konstantnom varijansom.

Međutim, od sedamdesetih godina dvadesetog veka, potpuno drugačiji pristup je imao sve veću primenu u analizi vremenskih serija. Čisto deskriptivne procedure zamenjene su rezultatima i metodama teorije verovatnoće i matematičke statistike. To je dovelo do drugačijeg posmatranja uloge stohastičkih kretanja u odnosu na vremenske serije. Pretpostavka modernog pristupa je da postoji stohastički uticaj na sve komponente vremenske serije. Dakle, 'zakon kretanja' cele vremenske serije odnosi se na stohastički proces, a analiza je samo jedna realizacija procesa generisanja podataka. Stohastički smisao sa uglavnom veoma kompleksnim zavisnim strukturama sada ima primarnu ulogu.

Prve korake u ovom pravcu napravili su ruski statističar *Evgeny Evgenievich Slutsky* i britanski statističar *George Udny Yule* početkom prošlog veka. *Yule* je još 1927. godine definisao modele vremenskih serija koji bi preciznije modelirali vremenske serije sa cikličnim fluktuacijama, a istovremeno poslužili za pouzdanije otkrivanje osobina tih fluktuacija. Njemu se pripisuje definisanje modela koji iskazuje zavisnost tekuće vrednosti od prethodnih i slučajne greške, tzv. *model autoregresionih procesa*. U isto vreme *Slutsky* (1927) je prvi definisao alternativni model vremenske serije koji iskazuje tekuću vrednost serije kao ponderisanu sumu slučajnih grešaka, tzv. *model pokretnih proseka*. U to vreme, u potrazi za cikličnim kretanjem u vremenskoj seriji, koristili su se postupci izravnjanja kojima se želelo doći do osnovne tendencije u seriji. *Herman Wold* (1938) je u svojoj doktorskoj disertaciji sistematizovao i generalizovao ove pristupe i povezao ih sa teorijom verovatnoće Kolmogorova.

Praktična upotreba ovih modela proširila se zahvaljujući George E.P. Box-u i Gwilym M. Jenkins-u (1970)<sup>2</sup>, koji su razvili metode za njihovu empirijsku implementaciju. Prepostavka je da postoji zajednički stohastički model za celu generaciju procesa vremenskih serija. Prvo, ovaj metod identificuje specifični model u osnovi određenih statističkih oblika. Drugo, parametri ovog modela su ocenjeni. Treće, specifikacija modela je proverena statističkim testovima. Ako greške specifikacije postanu očigledne, ona se mora izmeniti, a parametri opet oceniti. Ova procedura se ponavlja sve dok se ne konstruiše model koji zadovoljava zadate kriterijume. Takav model se, napisetku, može koristiti za prognozu.

Od skoro je ideja dekompozicije vremenskih serija ponovo zastupljena, naročito za modeliranje sezonskih varijacija. Međutim, nasuprot klasičnom pristupu, sada se prepostavlja da se sve komponente vremenskih serija mogu predstaviti jednostavnim stohastičkim modelima. Procedura desezoniranja vremenskih serija koju koristi *Eurostat*<sup>3</sup> (predstavništvo evropske statistike, smešteno u Luksemburgu) je takođe zasnovano na ovom pristupu.

Štaviše, od osamdesetih godina dvadesetog veka sve češće su se razmatrale potencijalne nestacionarnosti vremenskih serija. Eksplicitno su razmatrane pri konstruisanju modela, kad god je to bilo moguće i imalo smisla. Radovi posvećeni problemima ocenjivanja i testiranja tipa nestacionarnosti, tzv. testovi integrisanosti, poslednjih su godina dominirali u analizi ekonomskih vremenskih serija, a uporedo se ulažu napor i u rešavanju problema modeliranja višedimenzionalnih vremenskih serija, tzv. VAR modeli.

Decenija koja je za nama beleži najdinamičniji rast literature i referenci na temu ekonometrijske analize vremenskih serija. Kroz međusobno prožimanje ekonometrije i analize vremenskih serija nastoji se da se dođe do rezultata koji će predstavljati valjan okvir praktičnog modeliranja ekonomskih fenomena. [6, 7]

## 1.5 Ciljevi analize vremenskih serija

Praćenje neke pojave kroz vreme osnova je u nalaženju logike u sledu opservacije, otkrivanju uzroka promena, te u kontrolisanju i razumevanju same pojave i njenog toka. Kada postoji mogućnost da se objasni i predviđi naredni korak u numeričkoj realizaciji pojave i pri tom on ne odstupa suviše od prave vrednosti, govori se o kvalitetnoj analizi vremenske serije. U tom smislu, podrazumevaju se sledeća četiri osnovna cilja [7]:

- *Deskripcija* - prva etapa analize; obuhvata grafičke prikaze i sumarne statistike; izražava suštinske karakteristike posmatrane pojave.
- *Objašnjenje* - varijacija jedne serije može da posluži za objašnjenje varijacije druge, npr. regresiona analiza ili linearna transformacija serija; takođe, generalizacijom modela jednodimenzionalih dobijaju se modeli višedimenzionalih vremenskih serija.
- *Prognoza* - na osnovu prošlih opservacija identificuje se i oceni model (korišćenjem razuđenog skupa statističkih testova i kriterijuma za verifikovanje valjanosti modela), koji se potom koristi za formiranje prognoze budućih vrednosti serije.
- *Kontrola* - gradi se model funkcije prenosa vremenske serije i na osnovu njega se formira prognoza, a zatim se ulazna serija prilagođava tako da rezultujući proces bude što bliži željenom cilju.

2. George E.P. Box, Gwilym M. Jenkins (1970), *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco

3. <http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/>

## 1.6 Operatori u analizi vremenskih serija

**Definicija 1.6.1:** [7] Operator kašnjenja (docnje)<sup>4</sup>, u oznaci  $B$ , je algebarska veličina koja se definiše kao

$$BX_t = X_{t-1}, \quad (1.6.1)$$

a koja, jednostavno rečeno, tekuću opservaciju vremenske serije pomera jedan period nazad (u prošlost).

Prvo kašnjenje slučajne promenljive  $X_t$  u trenutku vremena  $t$  je  $X_{t-1}$ , drugo je  $X_{t-2}$ , itd. Kašnjenje dužine  $k$  predstavlja vrednost promenljive  $X$  koja je ostvarena  $k$  perioda pre trenutka vremena  $t$  i označava se kao  $X_{t-k}$ . Dakle, u opštem slučaju važi:

$$B^k X_t = X_{t-k}, \quad k \in \mathbb{Z}. \quad (1.6.2)$$

Definicija je kompletna tek kad se definiše i:  $B^0 X_t = X_t$ .

Očigledno da je za proizvoljnu konstantu  $c$ ,  $B^k c = c$ , tj. konstanta pomerena za  $k$  perioda unazad ima istu vrednost. Kada  $k$  uzima vrednosti iz skupa negativnih celih brojeva, tada je  $B^{-k} X_t = X_{t+k}$ , što znači da je tekuća opservacija pomerena  $k$  perioda unapred (u budućnost).

**Definicija 1.6.2:** [7] Operator prve diference definiše se na osnovu operatora kašnjenja i jediničnog operatora:

$$\Delta = \mathbf{1} - B. \quad (1.6.3)$$

Stoga je:

$$\Delta X_t = (\mathbf{1} - B)X_t = X_t - X_{t-1}, \quad (1.6.4)$$

što predstavlja razliku između vrednosti promenljive u trenutku  $t$  i trenutku  $t - 1$ .

Dvostrukom primenom operatora prve diference dobija se polinom drugog reda po operatoru kašnjenja, tj:

$$\Delta^2 = (1 - B)^2 = 1 - 2B + B^2 \quad (1.6.5)$$

Otuda je:

$$\begin{aligned} \Delta^2 X_t &= \Delta(\Delta X_t) = \Delta(X_t - X_{t-1}) = (X_t - X_{t-1}) - (X_{t-1} - X_{t-2}) = X_t - 2X_{t-1} + X_{t-2} \\ &= X_t - 2BX_t + B^2 X_t = (1 - 2B + B^2)X_t = (1 - B)^2 X_t. \end{aligned} \quad (1.6.6)$$

U opštem slučaju,  $\Delta^d = (1 - B)^d$  je polinom  $d$ -tog reda po operatoru kašnjenja, pa je

$$\Delta^d X_t = \Delta(\Delta^{d-1} X_t) = \dots = (1 - B)^d X_t. \quad (1.6.7)$$

---

4. Oznaka  $B$  je od prvog slova engleske reči *backward* (unazad). Na engleskom operator kašnjenja nosi naziv *lag operator*, pa se često u literaturi umesto oznake  $B$  koristi oznaka  $L$ .

## 2 STACIONARNE VREMENSKE SERIJE

### 2.1 Beli šum

**Definicija 2.1.1:** [7] Proces  $\{\varepsilon_t, t \in T\}$  naziva se *beli šum* (engl. *white noise*) ako su  $\varepsilon_1, \varepsilon_2, \dots$  nekorelisane slučajne veličine, nulte srednje vrednosti i stabilne varijanse, tj:

$$\bullet E(\varepsilon_t) = \mu_\varepsilon = 0, \quad t \in \mathbb{Z} \quad (2.1.1a)$$

$$\bullet \text{Var}(\varepsilon_t) = E(\varepsilon_t)^2 = \sigma_\varepsilon^2 = \text{const}, \quad t \in \mathbb{Z} \quad (2.1.1b)$$

$$\bullet \text{Cov}(\varepsilon_t, \varepsilon_{t-k}) = E(\varepsilon_t \varepsilon_{t-k}) = 0, \quad t \in \mathbb{Z}, k = 1, 2, \dots (k \neq 0). \quad (2.1.1c)$$

Autokovariansna i autokorelaciona funkcija belog šuma date su, respektivno, sa:

$$\gamma_k = \begin{cases} \sigma_\varepsilon^2 & , \quad k = 0 \\ 0 & , \quad k \neq 0 \end{cases} \quad \rho_k = \begin{cases} 1 & , \quad k = 0 \\ 0 & , \quad k \neq 0 \end{cases} \quad (2.1.2)$$

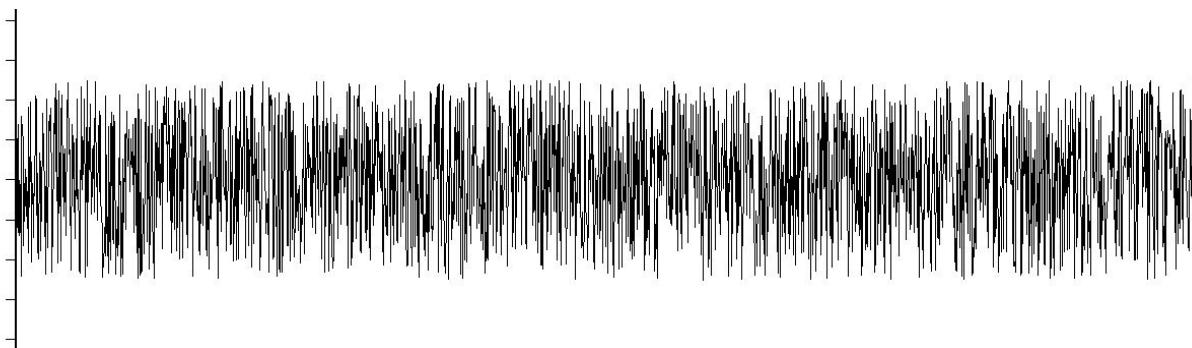
Iako se beli šum retko pojavljuje u praksi, on igra veoma važnu ulogu kao osnovni element u izgradnji modela vremenskih serija.

Ukoliko se navedenim uslovima u definiciji 2.1.1 doda i uslov da su članovi niza nezavisne slučajne veličine, koje su normalno raspodeljene, tada je razmatrani proces *Gausov beli šum*.

Tehnički posmatrano, nezavisni, jednako raspodeljeni proces (engl. *i.i.d.*) (sa konačnom varijansom) je beli šum, ali beli šum nije obavezno *i.i.d* proces (s obzirom na to da  $\varepsilon_t$  nisu obavezno identično raspodeljeni ili nezavisni) [17].

Beli šum je slučajni proces koji na izvestan način korespondira slučajnoj grešci klasičnog linearног regresionog modela. Sam termin izведен je iz spektralne analize bele svetlosti. Naime, spektar bele svetlosti karakteriše jednak doprinos svih sedam osnovnih boja spektra. Drugim rečima, ukupna energija bele svetlosti sadrži jednak uticaj komponenti, na različitim frekvencijama [10].

Na osnovu procesa belog šuma definiše se najopštija klasa procesa, tzv. *linearni procesi*.



Slika 2.1: Grafički prikaz belog šuma

## 2.2 Linearni procesi

Pri modeliranju, jedan od najbitnijih zadataka jeste pronalaženje odgovarajuće funkcionalne veze između određenih ulaznih i izlaznih informacija. *Linearan proces* predstavlja najopštiji okvir modeliranja stacionarnih vremenskih serija. Dakle, jedna od najčešćih prepostavki jeste da je veza između ulaznih i izlaznih parametara linearna. U analizi vremenskih serija linearni filter  $L$  je operator koji transformiše vremensku seriju  $\{X_t\}_{t \in \mathbb{Z}}$  u vremensku seriju  $\{Y_t\}_{t \in \mathbb{Z}}$ . [17, 12, 10, 3]

Vrednost vremenske serije  $\{Y_t\}_{t \in \mathbb{Z}}$  u trenutku  $t$  zavisi od sopstvenih prošlih i budućih vrednosti. Kako su u praksi na raspolaganju istorijski podaci, uvodi se prepostavka da je  $j \geq 0$ , pa se dolazi do definicije linearog filtera.

$$Y_t = L(X_t) = \sum_{j=0}^{\infty} \omega_j X_{t-j}, \quad (2.2.1)$$

gde vrednosti  $\{Y_t\}_{t \in \mathbb{Z}}$  u trenutku  $t$  zavise samo od sopstvenih prošlih vrednosti, a koeficijenti  $\omega_j$  su vremenski invarijantni. Ukoliko suma koeficijenata apsolutno konvergira, tj.  $\sum_{j=-\infty}^{\infty} |\omega_j| < \infty$ , kaže se da su stabilni.

**Definicija 2.2.1:** [3] Proces  $\{X_t, t \in T\}$  je *linearan* ako može da se predstavi u obliku

$$X_t = \mu_t + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}, \quad (2.2.2)$$

gde je  $\mu_t$  deterministička komponenta,  $\{\varepsilon_t, t \in T\}$  proces belog šuma, a  $\psi_1, \psi_2, \dots$  nepoznati parametri.

Osnovna karakteristika determinističke komponente je da ona može biti aproksimirana matematičkom funkcijom, što u situaciji prognoziranja vremenske serije znači da će izabrani tip funkcije važiti u budućem periodu. Suma  $\sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$  predstavlja stohastičku komponentu procesa i ona opisuje dejstvo slučajnih faktora u modelu. Osnovna prepostavka je da stohastička komponenta može da se modelira pomoću procesa belog šuma.

- Očekivanje i varijansa linearog procesa dati su sledećim izrazima [3]:

$$E(X_t) = E(\mu_t) + \sum_{j=0}^{\infty} \psi_j E(\varepsilon_{t-j}) = \mu_t + 0 = \mu_t \quad (2.2.3)$$

$$Var(X_t) = \sigma^2 \sum_{j=0}^{\infty} \psi_j^2 \quad (2.2.4)$$

- Momenti prvog i drugog reda su konačni u slučaju kada je  $\mu_t \equiv \mu = const$ ,  $\sum_{j=0}^{\infty} \psi_j^2 < \infty$ .

Tada su kovarijaciona i korelaciona funkcija:

$$\gamma_k = \sigma^2 \sum_{j=0}^{\infty} \psi_{j+k} \psi_j \quad (2.2.5)$$

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\sigma^2 \sum_{j=0}^{\infty} \psi_{j+k} \psi_j}{\sum_{j=0}^{\infty} \psi_j^2} \quad (2.2.6)$$

Ako su ispunjeni navedeni uslovi, kovarijaciona funkcija zavisi samo od  $k$ , pa je linearan proces slabo stacionaran. Ukoliko se dodatno pretpostavi da je beli šum Gausovski, tada je funkcija raspodele stohastičkog procesa potpuno određena prvim i drugim momentima procesa i proces je strogo stacionaran.

U slučaju kad važi  $\mu_t \equiv \mu = const$ , tada linearni proces ekvivalentno može da se zapiše u obliku:

$$X_t - \mu = X_t^* = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} \quad (2.2.7)$$

$$E(X_t^*) = 0$$

Godine 1938. Wold je dokazao teoremu koja se može smatrati egzistencijalnom. Smatrao je da je klasa linearnih procesa dovoljno opšta da obuhvati sve slabo stacionarne vremenske serije.

**TEOREMA 2.2.1:** (*Woldova dekompozicija*) [17] Stohastička komponenta slabo stacionarne vremenske serije  $\{X_t, t \in T\}$  može da se predstavi u formi linearog procesa oblika:

$$X_t - V_t = \sum_{j=1}^{\infty} X_j \varepsilon_{t-j}, \quad (2.2.8)$$

gde je  $V_t = \mu = E(X_t)$  deterministička komponenta,  $\{\varepsilon_t, t \in T\}$  je proces belog šuma,  $\sum_{j=1}^{\infty} X_j^2 < \infty$ ,  $X_0 = 1$ , a  $E(\varepsilon_t, V_t) = 0$ .

**Definicija 2.2.2:** [17] Linearan proces  $\{X_t, t \in T\}$  za koji važi da je  $E(X_t) = \mu$ ,  $\forall t \in T$  je *invertibilan* ako postoji  $\phi_1, \phi_2, \dots$  takvi da je:

$$\varepsilon_t = \sum_{j=0}^{\infty} \phi_j (X_{t-j} - \mu) = \sum_{j=0}^{\infty} \phi_j X_{t-j}^*, \quad \sum_{j=0}^{\infty} |\phi_j| < \infty. \quad (2.2.9)$$

**Definicija 2.2.3:** [10] *Linearna vremenska serija* je jedna realizacija linearog stohastičkog procesa.

Postoje tri grupe linearnih procesa koje imaju konačan broj parametara:

- *autoregresioni proces reda p (AR(p)),*
- *proces pokretnih proseka reda q (MA(q)),*
- *proces ARMA(p, q)* je kombinacija  $AR(p)$  i  $MA(q)$  procesa.

### 2.3 Autoregresioni procesi reda $p$ ( $AR(p)$ procesi)

Autoregresioni procesi, kao što im ime kazuje, impliciraju regresiju na sopstvene vrednosti, pa otuda i prefiks 'auto' u imenu ovog procesa. Imaju prirodnu interpretaciju - naredna opservacija je blago 'remećenje' jednostavne funkcije najskorijih opservacija. Korišćenjem standardnih softvera za regresiju lako je oceniti parametre i lako ih je prognozirati. [17, 10, 6, 30, 7]

**Definicija 2.3.1:** [17]  $\{X_t, t \in T\}$  je *autoregresioni proces reda  $p$*  ako se može predstaviti u obliku:

$$X_t = \mu + \sum_{j=1}^p \phi_j X_{t-j} + \varepsilon_t, \quad (2.3.1)$$

gde je  $\mu$  konstanta,  $\{\varepsilon_t, t \in T\}$  proces belog šuma, a  $\phi_j, j = 1, \dots, p$  nepoznati parametri.

S obzirom na to da je  $\mu$  konstanta, ona se može, sudeći po izrazu 2.2.7, izostaviti u zapisu  $AR(p)$  procesa:

$$X_t = \sum_{j=1}^p \phi_j X_{t-j} + \varepsilon_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t \quad (2.3.2)$$

S obzirom na to da je tekuća vrednost procesa linearna kombinacija  $p$  sopstvenih prošlih vrednosti plus slučajni poremećaj<sup>5</sup>  $\varepsilon_t$ , izraz 2.3.2 može se, korišćenjem operatora kašnjenja,  $B$ , sažetije predstaviti u sledećem obliku:

$$\phi(B)X_t = \varepsilon_t, \quad (2.3.3)$$

gde je  $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ . Kako je  $\sum_{j=1}^p |\phi_j| < \infty$ , proces je uvek invertibilan.

Jednačina  $\phi(z) = 0, z \in \mathbb{C}$  je karakteristična jednačina  $AR(p)$  procesa. Rešenja karakteristične jednačine su karakteristični korenji  $AR(p)$  procesa. Uslovi stacionarnosti autoregresionog procesa zavise od njih, a definisani su sledećom teoremom:

**TEOREMA 2.3.1:** [17] Neka su  $z_1, z_2, \dots, z_p$  karakteristični korenji autoregresionog procesa reda  $p$ . Ako za  $\forall z_i \in \mathbb{C}$  važi  $|z_j| > 1, j = 1, 2, \dots, p$ , odnosno  $|\phi_j| < 1$  (jer je  $z = \frac{1}{\phi} = \phi^{-1}$  karakteristični koren), onda je  $AR(p)$  proces *slabo stacionaran*.

Za korene karakteristične jednačine važe sledeće tvrdnje [10]:

- Ukoliko su svi korenji  $\phi_1, \phi_2, \dots, \phi_p$  po modulu strogo manji od jedan, onda je vremenska serija stacionarna.
- Ukoliko postoji bar jedan koren  $\phi_j$ , za  $j = 1, 2, \dots, p$ , koji je jednak vrednosti jedan po modulu, onda je vremenska serija nestacionarna. Takva vremenska serija se naziva *vremenska serija sa jediničnim korenom*. Ova forma nestacionarnosti se otklanja postupkom uzimanja razlike. Broj jediničnih korenova ovog tipa određuje koliko puta treba uzimati razlike vremenskih serija koje se tom prilikom dobijaju da bi vremenska serija postala stacionarna. Kada je bar jedno rešenje jednako minus jedan, tada u seriji postoji sezonski jedinični koren. Njegovo prisustvo otklanja se primenom sezonske diferencije.

5. Slučajni poremećaj obuhvata novine procesa u trenutku  $t$ , koje nisu objašnjene prošlim vrednostima serije.

- Ukoliko postoji bar jedan koren  $\phi_j$  za  $j = 1, 2, \dots, p$ , koji je strogo veći od jedan, tada je vremenska serija *eksplozivna*. To znači da je pod uticajem aditivnog dejstva trajno rastućeg efekta neočekivanih slučajnih šokova.

U slučaju kada je  $AR(p)$  slabo stacionaran, mogu se odrediti  $E(X_t)$ ,  $Var(X_t)$ ,  $\gamma_k$  i  $\rho_k$  [17]:

- Kako je  $E(X_t) = E(X_{t-1}) = \dots = E(X_{t-p})$ , iz definicije  $X_t$  sledi:

$$\begin{aligned} E(X_t) &= \sum_{j=1}^p \phi_j E(X_t) \quad \Rightarrow \quad (1 - \sum_{j=1}^p \phi_j) E(X_t) = 1 \\ &\Rightarrow \quad E(X_t) = \frac{1}{1 - \sum_{j=1}^p \phi_j} \end{aligned} \quad (2.3.4a)$$

- Autokovarijaciona funkcija ima sledeći oblik:

$$\begin{aligned} \gamma_k &= Cov(X_t, X_{t-k}) = Cov\left(\sum_{j=1}^p \phi_j X_{t-j} + \varepsilon_t, X_{t-k}\right) \\ &= \sum_{j=1}^p \phi_j Cov(X_{t-j}, X_{t-k}) + Cov(\varepsilon_t, X_{t-k}) \\ &= \sum_{j=1}^p \phi_j \gamma_{k-j} + \begin{cases} \sigma^2 & , \quad k = 0 \\ 0 & , \quad k > 0 \end{cases} \end{aligned} \quad (2.3.4b)$$

- Varijansa  $AR(p)$  procesa je:

$$\begin{aligned} Var(X_t) &= \gamma_0 = \sigma^2 + \sum_{j=1}^p \phi_j \gamma_{-j} = \sigma^2 + \sum_{j=1}^p \phi_j \gamma_j \\ &\Rightarrow \quad \gamma_0 - \sum_{j=1}^p \phi_j \gamma_j = \sigma^2 \\ &\Rightarrow \quad \gamma_0 \left(1 - \sum_{j=1}^p \phi_j \rho_j\right) = \sigma^2 \\ &\Rightarrow \quad Var(X_t) = \frac{\sigma^2}{1 - \phi_1 \rho_1 - \dots - \phi_p \rho_p} \end{aligned} \quad (2.3.4c)$$

- Autokorelaciona funkcija ima oblik:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p} \quad (2.3.4d)$$

### 2.3.1 Autoregresioni procesi prvog reda ( $AR(1)$ )

**Definicija 2.3.2:** [7]  $\{X_t, t \in T\}$  je autoregresioni proces prvog reda ako se može predstaviti u obliku:

$$X_t = \phi_1 X_{t-1} + \varepsilon_t. \quad (2.3.5)$$

Korišćenjem lag operatora  $B$  (operatora docnje/kašnjenja) formulacija 2.3.5 ima oblik:

$$(1 - \phi_1 B)X_t = \varepsilon_t, \quad (2.3.6)$$

odakle dobijamo:

$$\begin{aligned} X_t &= (1 - \phi_1 B)^{-1} \varepsilon_t = (1 + \phi_1 B + \phi_1^2 B^2 + \dots) \varepsilon_t \\ &= \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_1^2 \varepsilon_{t-2} + \dots = \sum_{k=0}^{\infty} \phi_1^k \varepsilon_{t-k}. \end{aligned} \quad (2.3.7)$$

S obzirom na to da su  $X_{t-1}$  i  $\varepsilon_t$  nekorelisani, sudeći po jednakosti 2.3.5 varijansa ovog procesa je:

$$Var(X_t) = \phi_1^2 Var(X_{t-1}) + \sigma_{\varepsilon}^2 \quad (2.3.8)$$

Ako je  $\{X_t, t \in T\}$  stacionaran, onda je  $Var(X_t) = Var(X_{t-1}) = \sigma_x^2$ , pa važi:

$$\sigma_x^2 = \phi_1^2 \sigma_x^2 + \sigma_{\varepsilon}^2. \quad (2.3.9)$$

To dalje implicira da je  $\sigma_x^2 > \phi_1^2 \sigma_x^2$ , tj.

$$1 > \phi_1^2. \quad (2.3.10)$$

Ako je  $|\phi_1| < 1$  ovaj proces konvergira u srednje kvadratnom, jer je  $\sum |\phi_1|^{2k} < \infty$ . Granični proces je stacionaran i zadovoljava jednačinu 2.3.5. Dakle, ako je  $|\phi_1| < 1$ , onda postoji stacionarno rešenje jednačine 2.3.5. To je uslov stacionarnosti AR procesa prvog reda.

Ako je  $|\phi_1| > 1$  proces definisan sa  $(1 - \phi_1 B)^{-1} = \sum_{k=0}^{\infty} \phi_1^k B^k$  nije konvergentan. Međutim, formulacija 2.3.5 može se preformulisati u oblik:

$$X_t = \frac{1}{\phi_1} X_{t+1} - \frac{1}{\phi_1} \varepsilon_{t+1}, \quad (2.3.11)$$

pa korišćenjem istog argumenta dobijamo i:

$$X_t = - \sum_{k=0}^{\infty} \phi_1^{-k} \varepsilon_{t+k}. \quad (2.3.12)$$

Ovaj proces jeste konvergentan, pa postoji stacionarna serija  $X_t$  koja zadovoljava 2.3.5. Rezultujući proces u opštem slučaju nije zadovoljavajući za modeliranje ili prognozu, jer nije uzročni (zavisi od budućih vrednosti  $\varepsilon_t$ ).

Ako je  $|\phi_1| = 1$  ne postoji stacionarno rešenje za 2.3.5.

Sudeći po svemu prethodnom, u cilju modeliranja ili prognoze stacionarnih vremenskih serija, pažnja se mora ograničiti na serije koje imaju  $|\phi_1| < 1$ , tj. na one koje zadovoljavaju uslov stacionarnosti. Što se invertibilnosti tiče, i za ove procese, kao i za uopšteno  $AR$  procese konačnog reda, važi da su svi invertibilni [3].

Ako se obe strane jednačine 2.3.5 pomnože sa  $X_{t-k}$ ,  $k \geq 0$  i prođe se očekivanjem, dobija se: [3, 7]

$$E(X_t X_{t-k}) = \phi_1 E(X_{t-1} X_{t-k}) + E(\varepsilon_t X_{t-k}). \quad (2.3.13)$$

Drugi sabirak s desne strane jednak je nuli, jer je, zbog linearne reprezentacije procesa,  $\varepsilon_t$  nezavisno od ranijih vrednosti  $X_t$ . To znači da autokovarijacija mora da zadovoljava rekurziju  $\gamma_k = \phi_1 \gamma_{k-1}$ ,  $k = 1, 2, 3, \dots$ . Ovo je linearna diferencijalna jednačina prvog reda sa rešenjem  $\gamma_k = \phi_1^k \gamma_0$ ,  $k = 0, 1, 2, \dots$

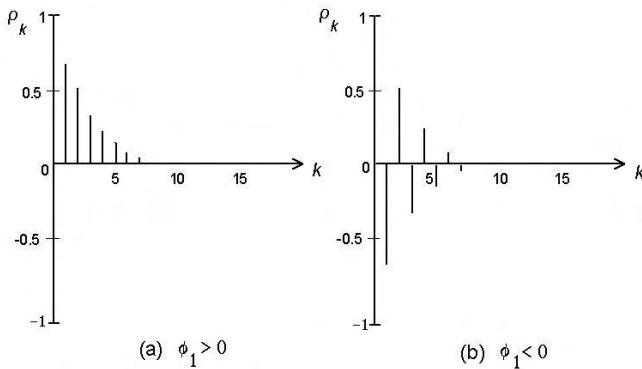
Preuređivanjem jednačine 2.3.9 dobija se da je  $\gamma_0 = \sigma_\varepsilon^2 / (1 - \phi_1^2)$ , pa odatle sledi:

$$\gamma_k = \frac{\phi_1^k \sigma_\varepsilon^2}{1 - \phi_1^2}, \quad k = 0, 1, 2, \dots \quad (2.3.14)$$

To, pak, znači da je autokorelaciona funkcija ( $\rho_k / \gamma_0$ ) data sa:

$$\rho_k = \phi_1^k, \quad k = 0, 1, 2, \dots \quad (2.3.15)$$

Ako je  $\phi_1 > 0$  vrednosti autokorelace funkcije opadaju po eksponencijalnoj krivoj ka nuli, dok za  $\phi_1 < 0$  vrednosti autokorelace funkcije alterniraju (često menjaju smer), opadajući ka nuli, odnosno predstavljaju kombinaciju dve opadajuće eksponencijalne krive. U oba slučaja opadanje je sporije ukoliko je  $\phi_1$  blisko granicama stacionarnosti: +1 ili -1.



Slika 2.2: Autokorelaciona funkcija  $AR(1)$  modela, [7]

### 2.3.2 Autoregresioni procesi drugog reda ( $AR(2)$ )

**Definicija 2.3.3:** [7] Autoregresioni proces drugog reda, u oznaci  $AR(2)$ , definisan je sa:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t, \quad (2.3.16)$$

ili, korišćenjem operatora:

$$(1 - \phi_1 B - \phi_2 B^2) X_t = \varepsilon_t. \quad (2.3.17)$$

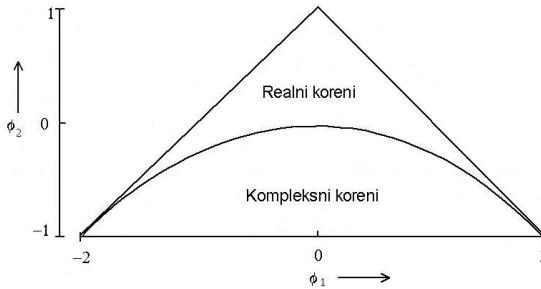
Kao i kod  $AR(1)$  modela, i ovde se može razmatrati invertibilnost  $AR$  operatora. Da bi se proverilo da li je uopšte moguća, faktoriše se operator  $1 - \phi_1 B - \phi_2 B^2$  i invertuje svaki faktor pojedinačno. Dakle, prepostavi se da je:

$$1 - \phi_1 B - \phi_2 B^2 = (1 - c_1 B)(1 - c_2 B). \quad (2.3.18)$$

Invertovanje svakog od faktora odvojeno moguće je ako je  $|c_1| < 1$  i  $|c_2| < 1$ , ili ekvivalentno, ako su korenji polinoma  $1 - \phi_1 z - \phi_2 z^2$  van jediničnog kruga. Uz malo algebre, dobija se da je ovo ekvivalentno sledećim uslovima:

$$\phi_1 + \phi_2 < 1, \quad -\phi_1 + \phi_2 < 1, \quad -1 < \phi_2 < 1. \quad (2.3.19)$$

Ovi uslovi se nazivaju *uslovi stacionarnosti AR(2) procesa*.



Slika 2.3: Oblast stacionarnih rešenja AR(2) procesa, [7]

Ako se obe strane jednakosti 2.3.16 pomnože sa  $X_{t-k}$  i pusti očekivanje dobija se:

$$E(X_t X_{t-k}) = \phi_1 E(X_{t-1} X_{t-k}) + \phi_2 E(X_{t-2} X_{t-k}) + E(\varepsilon_t X_{t-k}), \quad (2.3.20)$$

odakle sledi rekurzija:

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2}, \quad (2.3.21)$$

sa početnim uslovima

$$\begin{aligned} \gamma_0 &= \phi_1 \gamma_1 + \phi_2 \gamma_2 + \sigma_\varepsilon^2 \\ \gamma_1 &= \phi_1 \gamma_0 + \phi_2 \gamma_1. \end{aligned} \quad (2.3.22)$$

Podelom izraza 2.3.21 sa  $\gamma_0$  dobija se autokorelaciona funkcija:

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2}, \quad k \geq 1. \quad (2.3.23)$$

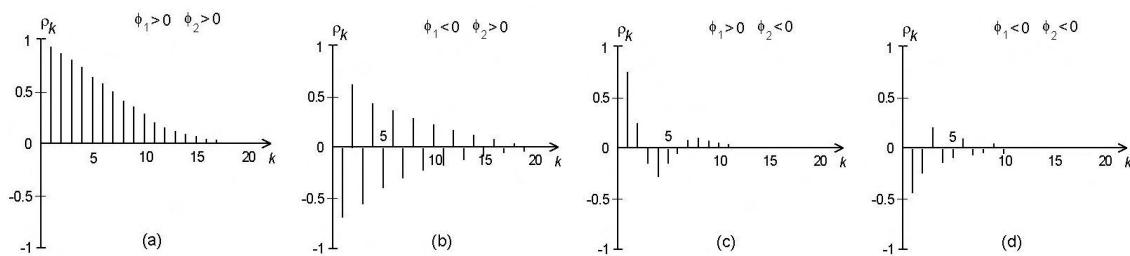
Rešenje ove diferencijalne jednačine može se napisati u obliku:

$$\rho_k = A_1(z_1^{-1})^k + A_2(z_2^{-1})^k, \quad k = 0, 1, 2, \dots, \quad (2.3.24)$$

gde su  $z_1$  i  $z_2$  korenji polinoma  $1 - \phi_1 z - \phi_2 z^2$ , a  $A_1$  i  $A_2$  konstante čije vrednosti zavise od početnih vrednosti  $\rho_0 = 1$  i  $\rho_{-1} = \rho_1$ . U slučaju kad su korenji jednakci, rešenje ima opšti oblik:

$$\rho_k = (A_1 + A_2 k) z^{-k} \quad (2.3.25)$$

Ova jednačina pokazuje da će autokorelaciona funkcija AR(2) serija sa jednakim korenima eksponencijalno opadati kako  $k \rightarrow \infty$  [3,7].



Slika 2.4: Autokorelacione funkcije AR(2) procesa, [7]

Na slici 2.4 prikazane su autokorelacione funkcije  $AR(2)$  procesa za četiri karakteristične kombinacije koeficijenata  $\phi_1$  i  $\phi_2$ . Kada su rešenja karakteristične jednačine realna (slučaj (a) i (b)) autokorelacione funkcije su kombinacija dve prigušene eksponencijalne funkcije, a za kompleksna rešenja (slučaj (c) i (d)) autokorelacione funkcije slede putanju prigušene sinusne funkcije.

## 2.4 Procesi pokretnih proseka reda $q$ ( $MA(q)$ procesi)

Proces pokretnih proseka (engl. *moving average*) je koristan u modeliranju pojava kod kojih događaji uzrokuju trenutne efekte, a koji traju kratak period vremena. Ovaj proces se dovodi u vezu sa postupkom izravnjanja vremenske serije. Pokretne sredine se mogu računati na osnovu aritmetičke sredine i medijane. Za izravnanje vremenske serije na osnovu aritmetičke sredine postupak je sledeći:

Neka je dat niz podataka:

$$X_1, X_2, \dots, X_{t-m}, \dots, X_{t-2}, X_{t-1}, X_t, X_{t+1}, X_{t+2}, \dots, X_{t+m}, \dots, X_T$$

Svaki podatak  $X_t$  zamenjuje se novim podatkom  $Y_t$  dobijenim na sledeći način:

$$Y_t = \sum_{i=-m}^m a_i X_{t-i}, \quad t = m+1, \dots, T-m, \quad \sum_{i=-m}^m a_i = 1.$$

Navedeni postupak izravnjanja se može primeniti više puta, čime se gubi prvih  $m$  i poslednjih  $m$  podataka. Izravnjanjem se dobija vremenska serija približno iste aritmetičke sredine, ali sa znatno manjim varijabilitetom [10].

**Definicija 2.4.1:** [7]  $\{X_t, t \in T\}$  je proces pokretnih proseka reda  $q$  ako se može predstaviti u obliku:

$$X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}, \quad (2.4.1)$$

koji se, korišćenjem *lag* operatora, sažetije predstavlja na sledeći način:

$$X_t = \theta(B)\varepsilon_t, \quad (2.4.2)$$

gde je  $\theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$ , a  $\{\varepsilon_t, t \in T\}$  beli šum.

U datom modelu nivo vremenske serije u trenutku  $t$  opisuje se u funkciji od članova procesa belog šuma u trenucima  $t, t-1, \dots, t-q$ . Pomeranjem u vremenu menjaju se članovi zbiru, ali broj sabiraka u zbiru ostaje isti. Proces pokretnih proseka reda  $q$  je linearni proces sa konačnim očekivanjem i konačnim brojem parametara.

Autokovarijaciona funkcija  $MA(q)$  procesa je [30]:

$$\gamma_k = \begin{cases} (1 + \theta_1^2 + \dots + \theta_q^2)\sigma^2 & , \quad k = 0 \\ (-\theta_k + \theta_1\theta_{k+1} + \dots + \theta_{q-k}\theta_q)\sigma^2 & , \quad k = 1, \dots, q \\ 0 & , \quad k > q \end{cases} \quad (2.4.3)$$

a autokorelaciona funkcija:

$$\rho_k = \begin{cases} \frac{-\theta_k + \theta_1\theta_{k+1} + \dots + \theta_{q-k}\theta_q}{1 + \theta_1^2 + \dots + \theta_q^2} & , \quad k = 1, 2, \dots, q \\ 0 & , \quad k > q. \end{cases} \quad (2.4.4)$$

Dakle, pošto za proces pokretnih proseka važi da je  $E(X_t) < \infty$ ,  $Var(X_t) < \infty$  i  $\gamma_k = f(k)$ , očigledno je da je dati proces uvek slabo stacionaran [17].

Sudeći po ponašanju autokorelace funkcije, uvek je lako uvideti razliku između  $MA$  i  $AR$  serija: za  $MA$  serije su karakteristični oštiri prekidi, dok autokorelace funkcije  $AR$  procesa opadaju eksponencijalno [7].

**TEOREMA 2.4.1:** [18] Proces pokretnih proseka reda  $q$  je *invertibilan* ako su svi korenii njegove karakteristične jednačine po modulu veći od jedinice ili, ekvivalentno, ako su svi parametri  $\theta_i$ ,  $i = 1, 2, \dots, q$  po modulu manji od jedinice.

Ukoliko je proces invertibilan kaže se da on ima  $AR(\infty)$  reprezentaciju. Neka su dati  $AR(p)$  i  $MA(q)$  proces respektivno:  $\varepsilon_t = \phi(B)X_t$  i  $X_t = \theta(B)\varepsilon_t$ . Tada je:

$$X_t = \theta(B)\varepsilon_t = \theta(B)\phi(B)X_t \quad \Rightarrow \quad \theta(B)\phi(B) = 1 \quad \Rightarrow \quad \theta(B) = \phi(B)^{-1}, \quad (2.4.5)$$

što daje vezu između koeficijenata procesa.

Stacionarnom autoregresionom procesu konačnog reda odgovara proces pokretnih proseka beskonačnog reda i invertibilnom procesu pokretnih proseka konačnog reda odgovara autoregresioni proces beskonačnog reda. Uzajamna veza između koeficijenata ova dva modela našla je odraza u odnosu njihovih korelacionih i parcijalnih korelacionih koeficijenata. Korelaciona funkcija  $AR(p)$  procesa i parcijalna korelaciona funkcija  $MA(q)$  procesa lagano odumiru ka 0, a parcijalna korelaciona funkcija  $AR(p)$  i korelaciona funkcija  $MA(q)$  su jednake 0 kada je  $k$  veće od reda procesa [7].

#### 2.4.1 Procesi pokretnih proseka prvog reda ( $MA(1)$ )

Kada se u izraz 2.4.1 uvrsti  $q = 1$  dobija se pokretni prosek prvog reda.

**Definicija 2.4.2:** [7]  $\{X_t, t \in T\}$  je *proces pokretnih proseka prvog reda* ako se može predstaviti u obliku:

$$X_t = \varepsilon_t - \theta_1\varepsilon_{t-1} = (1 - \theta_1B)\varepsilon_t \quad (2.4.6)$$

Za  $MA(1)$  proces matematičko očekivanje je, očigledno, nula ( $E(X_t) = 0$ ).

Autokovarijaciona i autokorelaciona funkcija ovog procesa su respektivno date izrazima:

$$\gamma_k = \begin{cases} (1 + \theta_1^2)\sigma^2 & , \quad k = 0 \\ -\theta_1\sigma^2 & , \quad k = 1 \\ 0 & , \quad k > 1 \end{cases} \quad \rho_k = \begin{cases} \frac{-\theta_1}{1 + \theta_1^2} & , \quad k = 1 \\ 0 & , \quad k > 1. \end{cases} \quad (2.4.7)$$

Odavde sledi da je tekuća vrednost procesa korelisana samo sa prethodnom i narednom vrednošću, što sugerije da je 'memorija' procesa ograničena samo na jedan period.

Za ma koju vrednost koeficijenta  $\theta_1$ ,  $MA(1)$  proces  $X_t = (1 - \theta_1B)\varepsilon_t$  i  $X_t = (1 - \frac{1}{\theta_1}B)\varepsilon_t$  imaće istu autokorelacionu funkciju, što znači da se na osnovu autokorelace funkcije (koja je kod  $MA(1)$  procesa  $-0.5 < \rho_k < 0.5$ , za svako  $k$ ) ne može jednoznačno identifikovati  $MA$  proces [7].

U nekim slučajevima poželjno je  $MA$  proces predstaviti u formi  $AR$  reprezentacije. U tom cilju  $MA(1)$  proces pišemo u obliku:

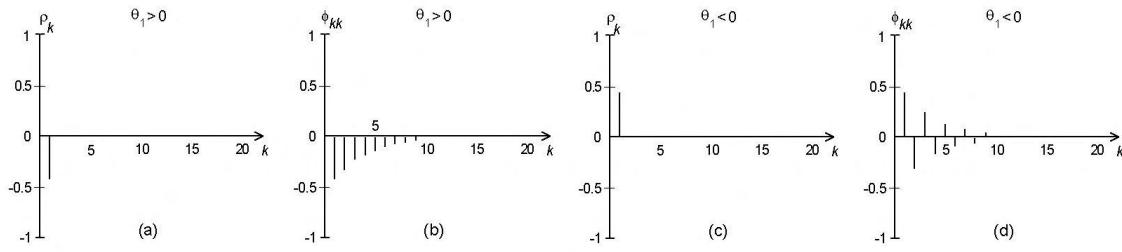
$$\begin{aligned}\varepsilon_t &= X_t + \theta_1 \varepsilon_{t-1} \\ &= X_t + \theta_1(X_{t-1} + \theta_1 \varepsilon_{t-2}) = X_t + \theta_1 X_{t-1} + \theta_1^2 \varepsilon_{t-2} \\ &= X_t + \theta_1 X_{t-1} + \theta_1^2(X_{t-2} + \theta_1 \varepsilon_{t-3}) = \dots \\ &= X_t + \theta_1 X_{t-1} + \theta_1^2 X_{t-2} + \theta_1^3 X_{t-3} + \dots\end{aligned}\quad (2.4.8)$$

odnosno

$$X_t = (-\theta_1 X_{t-1} - \theta_1^2 X_{t-2} - \dots) + \varepsilon_t. \quad (2.4.9)$$

Uz uslov  $|\theta_1| < 1$  dobija se konvergentna autoregresiona reprezentacija  $MA(1)$  procesa. Dakle, ako je  $|\theta_1| < 1$ ,  $MA(1)$  proces se može izraziti kao  $AR(\infty)$  proces. U takvom slučaju se kaže da je  $MA$  proces *invertibilan*, a navedeni uslov se naziva *uslov invertibilnosti* [3].

Nasuprot autokorelacionoj funkciji koja se prekida posle prve docnje, parcijalna autokorelaciona funkcija za  $MA(1)$  model lagano odumire na dva načina, u zavisnosti od predznaka  $\theta_1$ , odnosno predznaka  $\rho$ . Ako je  $\theta_1$  negativnog predznaka (slika 2.5 d)), vrednost parcijalne autokorelacione funkcije na prvoj docnji je pozitivna da bi zatim alternirala; u suprotnom opada, uzimajući samo negativne vrednosti (slika 2.5 pod a)) [7].



Slika 2.5: Obična i parcijalna autokorelaciona funkcija  $MA(1)$  procesa, [7]

## 2.4.2 Procesi pokretnih proseka drugog reda ( $MA(2)$ )

Ako se u izrazu 2.4.1 zameni  $q = 2$ , dobija se proces pokretnog proseka drugog reda.

**Definicija 2.4.3:** [7]  $\{X_t, t \in T\}$  je *proces pokretnih proseka drugog reda* ako se može predstaviti u obliku:

$$X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} = (1 - \theta_1 B - \theta_2 B^2) \varepsilon_t \quad (2.4.10)$$

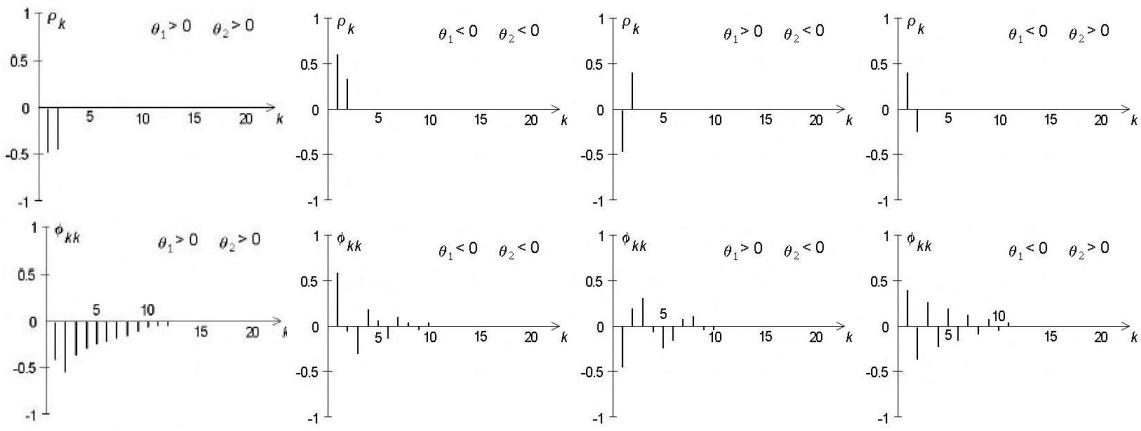
$MA(2)$  proces je, kao i svi procesi pokretnih proseka, stacionaran. Uslov invertibilnosti ispunjen je kada su korenji karakteristične jednačine ovog procesa po modulu manji od jedinice, odnosno, ekvivalentno tome, kada je:

$$\theta_1 + \theta_2 < 1, \quad \theta_2 - \theta_1 < 1, \quad -1 < \theta_2 < 1. \quad (2.4.11)$$

Autokovarijacioni i autokorelacioni koeficijenti, redom, iznose:

$$\gamma_k = \begin{cases} (1 + \theta_1^2 + \theta_2^2)\sigma^2 & , \quad k = 0 \\ (-\theta_1 + \theta_1 \theta_2)\sigma^2 & , \quad k = 1 \\ -\theta_2 \sigma^2 & , \quad k = 2 \\ 0 & , \quad k > 2 \end{cases} \quad \rho_k = \begin{cases} \frac{-\theta_1(1 - \theta_2)}{1 + \theta_1^2 + \theta_2^2} & , \quad k = 1 \\ \frac{-\theta_2}{1 + \theta_1^2 + \theta_2^2} & , \quad k = 2 \\ 0 & , \quad k > 2. \end{cases} \quad (2.4.12)$$

Model implicira da su opservacije međusobno nekorelisane ako su udaljene više od dva perioda jedna od druge. [3, 7]



Slika 2.6: Obična i parcijalna autokorelaciona funkcija  $MA(2)$  procesa, [7]

Na slici 2.6 prikazane su obična i parcijalna autokorelaciona funkcija svih karakterističnih slučajeva  $MA(2)$  procesa. Vrednosti ovih funkcija eksponencijalno opadaju ili osciliraju sledeći prigušenu sinusoidu u zavisnosti od predznaka i veličine koeficijenata  $\theta_1$  i  $\theta_2$ .

## 2.5 Autoregresioni procesi pokretnih proseka ( $ARMA(p, q)$ procesi)

Kao što je do sada pokazano, stacionaran i invertibilan proces se može prikazati ili u autoregresionej formi ili formi pokretnih proseka. Problem sa bilo kojom reprezentacijom je taj što ona može sadržati preveliki broj parametara, čak i u slučaju  $MA$ , odnosno  $AR$  modela konačnog reda, jer su za dobru aproksimaciju često potrebni modeli višeg reda. Uopšte, veliki broj parametara umanjuje efikasnost ocenjivanja. Stoga, u modeliranju može biti neophodno uključiti i autoregresione komponente i komponente pokretnih proseka, što nas dovodi do mešovitog autoregresionog procesa pokretnih proseka ( $ARMA$  procesa). [17, 10, 7, 3, 30]

**Definicija 2.5.1:** [17]  $\{X_t, t \in T\}$  je  $ARMA(p, q)$  proces ukoliko se može predstaviti u obliku:

$$X_t = \sum_{j=1}^p \phi_j X_{t-j} + \varepsilon_t - \sum_{i=1}^q \theta_i \varepsilon_{t-i}, \quad (2.5.1)$$

ili u ekvivalentnoj formi:

$$\phi(B)X_t = \theta(B)\varepsilon_t, \quad (2.5.2)$$

gde je:

- $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$  AR karakteristični polinom  $ARMA(p, q)$  procesa,
- $\theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$  MA karakteristični polinom  $ARMA(p, q)$  procesa.

U  $ARMA(p, q)$  modelu, jasno,  $p$  predstavlja red autoregresione komponente, a  $q$  red komponente pokretnih proseka. Specijalni slučajevi  $ARMA(p, q)$  procesa su  $AR(p) = ARMA(p, 0)$  i  $MA(q) = ARMA(0, q)$ . Drugačije se još  $ARMA(p, q)$  proces može zapisati:

$$X_t - \underbrace{\sum_{j=1}^p \phi_j X_{t-j}}_{AR \text{ komponenta}} = \varepsilon_t - \underbrace{\sum_{i=1}^q \theta_i \varepsilon_{t-i}}_{MA \text{ komponenta}} \quad (2.5.3)$$

Kako je proces pokretnih sredina konačnog reda uvek stacionaran i autoregresioni proces konačnog reda uvek invertibilan, onda važi sledeće:

- $ARMA(p, q)$  proces je stacionaran ako  $AR$  komponenta ispunjava uslove stacionarnosti  
 $\Leftrightarrow \forall z \in \mathbb{C}$  za koje je  $\phi(z) = 0$  važi  $|z| > 1$ ,
- $ARMA(p, q)$  proces je invertibilan ako  $MA$  komponenta ispunjava uslove invertibilnosti  
 $\Leftrightarrow \forall z \in \mathbb{C}$  za koje je  $\theta(z) = 0$  važi  $|z| > 1$ .

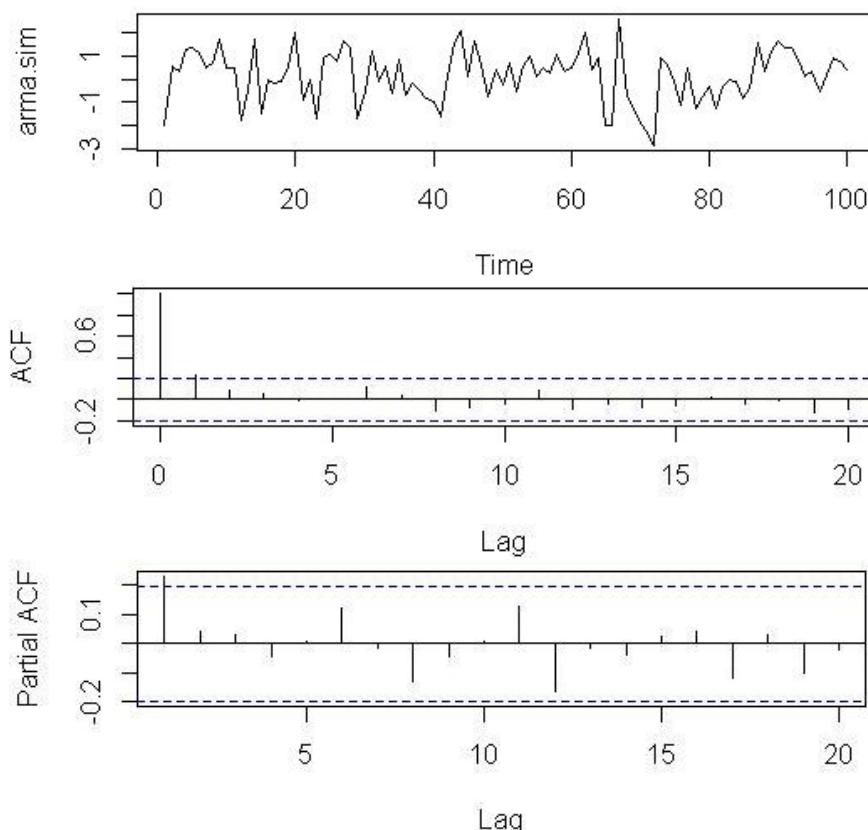
Ukoliko je  $ARMA(p, q)$  proces stacionaran, onda se može predstaviti kao proces pokretnih proseka sa beskonačno mnogo parametara, a ukoliko je invertibilan može se predstaviti kao autoregresioni proces sa beskonačno mnogo parametara.

Ako se izraz 2.5.1 raspšire u obliku  $X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$ , množenjem leve i desne strane sa  $X_{t-k}$  i određivanjem očekivane vrednosti tog izraza dobijaju se autokovarijaciona i autokorelaciona funkcija  $ARMA(p, q)$  procesa:

$$\gamma_k = \phi_1 \gamma_{k-1} + \dots + \phi_p \gamma_{k-p}, \quad k \geq q+1 \quad i \quad \rho_k = \phi_1 \rho_{k-1} + \dots + \phi_p \rho_{k-p}, \quad k \geq q+1. \quad (2.5.4)$$

Kao kod  $AR(p)$  procesa i autokorelaciona funkcija  $ARMA(p, q)$  procesa 'odumire' ka nuli nakon  $q$  koraka, što zavisi od autoregresionih koeficijenata procesa. Međutim, bitna razlika je u tome što prvih  $q$  autokorelacionih koeficijenata  $ARMA$  procesa,  $\rho_q, \rho_{q-1}, \dots, \rho_1$ , zavise kako od autoregresionih tako i od koeficijenata pokretnih proseka. Slično je sa parcijalnom autokorelacionom funkcijom koja je pod uticajem  $MA$  dela procesa [7].

Programom u statističkom softveru  $R$  može se simulirati npr. 100 observacija  $ARMA(2, 2)$  procesa i na osnovu toga iscrtati grafik serije, kao i autokorelacione i parcijalne autokorelacione funkcije (slika 2.7).



Slika 2.7: Simulacija  $ARMA(2,2)$  procesa u R-u

### 3 NESTACIONARNE VREMENSKE SERIJE

U prethodnoj glavi istaknute su osnovne karakteristike stacionarnih stohastičkih procesa: konstantnost sredine, tj. nivoa serije, konstantnost varijanse i zavisnost autokovarijansi samo od vremenskih intervala. Međutim, većina vremenskih serija, naročito u ekonomiji i poslovanju, ima bar jednu osobinu kojom odstupa od karakterističnog izgleda stacionarne serije. [7, 10, 30]

Stohastički procesi i vremenske serije sa vremenski zavisnim nivoom i/ili varijansom nazivaju se *nestacionarni stohastički procesi i serije*. Dakle, oni mogu imati ili nekonstantnu srednju vrednost ili vremenski promenljive momente drugog reda, ili i jedno i drugo. Stoga, razlikuju se procesi koji su *nestacionarni u srednjem* (*u odnosu na srednju vrednost*) i oni koji su *nestacionarni u odnosu na varijansu*.

Klasa nestacionarnih procesa obuhvata dve grupe:

- procesi koji imaju eksplozivan tok,
- procesi koji transformacijama mogu da se svedu na stacionarne linearne procese.

Kod nestacionarnih procesa koji imaju eksplozivan tok varijansa neograničeno raste u toku vremena, pa ova grupa nestacionarnih linearnih procesa nema praktičnog značaja za analizu vremenskih serija. Druga grupa obuhvata sve one procese kod kojih se stacionarnost postiže odgovarajućim transformacijama. Da bi se definisale transformacije potrebno je utvrditi uzrok nestacionarnosti procesa.

#### 3.1 Stabilizacija nestacionarnosti u odnosu na srednju vrednost

U klasičnoj analizi vremenskih serija, uobičajeni način modeliranja nestacionarnosti u odnosu na srednju vrednost polazi od toga da se srednja vrednost nestacionarnog procesa može predstaviti determinističkom funkcijom vremena. Tada se za seriju kaže da joj odgovara *deterministički trend* [7]. Postojanje trenda označava dugoročnu komponentu u kretanju ekonomske vremenske serije. U opštem slučaju, pretpostavka je da sredina evoluira kao polinom po vremenu reda  $d$ . Ocenom izabrane linije trenda i računanjem serije reziduala dobija se stacionaran stohastički proces. Ovo je klasa *trend stacionarnih procesa* i predstavlja se kao zbir determinističke komponente  $\mu_t$  (polinom po vremenu) i stacionarne komponente  $\eta_t$ , tj.:

$$X_t = \mu_t + \eta_t = \sum_{j=0}^d \beta_j t^j + \phi(B) \varepsilon_t. \quad (3.1.1)$$

Pošto je  $E(\eta_t) = \phi(B)E(\varepsilon_t) = 0$  i  $E(X_t) = E(\mu_t) = \sum_{j=0}^d \beta_j t^j$ , a koeficijenti  $\beta_j$  ostaju konstantni tokom vremena, ovaj trend je deterministički. Ključna pretpostavka od koje polaze ovi modeli, pogotovo u situaciji njihovog korišćenja u predviđanju, je da će izabrani tip funkcije trenda važiti i u budućem periodu [7].

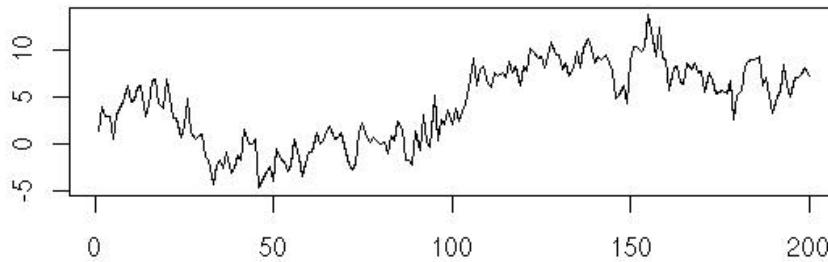
Veoma često se u primenama javlja situacija da se za vremensku seriju ne može prihvatiti pretpostavka postojanja determinističke funkcije vremena kojom bismo u celom posmatranom periodu modelirali tu seriju. Tako na primer, za seriju koja ima sve karakteristike nestacionarne

serije, ali se istovremeno za nju ne može pretpostaviti postojanje globalnog determinističkog trenda, i sam trend doživljava promene tokom vremena pod dejstvom slučajnih faktora. Takav trend se naziva *stohastički trend*. On se menja pod uticajem slučajnih faktora i označava dugoročnu tendenciju razvoja, koja se ne može predvideti na osnovu poznavanja podataka u prošlosti. Ovaj način modeliranja zasnovan je na pretpostavci da će diferenca reda  $d$  nestacionarne serije u srednjem biti stacionarna. Da bi se bolje ilustroval postupak transformacije potrebno je definisati proces pod nazivom *slučajni hod*.

**Definicija 3.1.1:** [10]  $\{X_t, t \in T\}$  je proces *slučajnog hoda* ako važi:

$$X_t = X_{t-1} + \varepsilon_t, \quad (3.1.2)$$

gde je  $\{\varepsilon_t, t \in T\}$  proces belog šuma. Naziv slučajni hod potiče od karakterističnog ponašanja koje podseća na putanju pijanog čoveka koji slučajno krivuda dok se kreće. Vremensku seriju koja prati slučajan hod karakterišu periodi rastućeg trenda, gde se trend menja iznenada i promena je nepredvidiva.



Slika 3.1: Grafik slučajnog hoda

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

$$X_t = X_0 + \sum_{i=1}^t \varepsilon_i, \quad (3.1.3)$$

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

$$X_t = X_{t-1} + \varepsilon_t / - X_{t-1} \Rightarrow \Delta X_t = \varepsilon_t \quad (3.1.4)$$

Operacija se zove diferenciranje, a operator  $\Delta$  diferencni operator. Procesi kod kojih se stacionarnost postiže diferenciranjem nazivaju se *diferencno stacionarni procesi*. Nekada je potrebno više puta diferencirati proces da bi se postigla stacionarnost [10].

Iako oba trenda (i deterministički i stohastički) opisuju tendenciju dugoročne promene, u pitanju su dva potpuno različita pristupa u modeliranju vremenske serije. Razlika se uočava u zavisnosti od toga da li se promene serije tokom vremena mogu predvideti ili ne.

Ako je, na primer, za određenu vremensku seriju utvrđeno da je trend stacionarna serija, to znači da se beleži rast po konstantnoj stopi koja je implicirana koeficijentima polinoma  $\mu_t$ . S druge strane, ako je serija diferencno stacionarna, tada i kratkoročni poremećaji imaju odraza na dugoročni rast. Otuda interes analitičara da se ispita karakter serije.

### 3.2 Stabilizacija nestacionarnosti u odnosu na varijansu

Proces nestacionaran u odnosu na očekivanje biće takođe nestacionaran i u odnosu na varijansu, ali proces stacionaran u odnosu na očekivanje nije obavezno stacionaran i u odnosu na varijansu. Postupak diferenciranja nije odgovarajući postupak transformacije, ako je posmatrani proces nestacionaran samo u varijansi. Tada je zadatak, zbog prepostavljene konstantnosti varijanse u teoriji stacionarnih stohastičkih procesa, naći transformaciju kojom će se varijansa stabilizovati. [20, 7]

Neka se varijansa nestacionarnog procesa menja sa promenom sredine (nivoa) serije, tj:

$$\text{Var}(X_t) = cf(\mu_t), \quad (3.2.1)$$

za neku pozitivnu konstantu  $c$ , funkciju  $f$  i očekivanu vrednost  $\mu_t$ .

U cilju nalaženja funkcije  $T$ , tako da transformisana serija  $T(X_t)$  ima konstantnu varijansu, željenu funkciju treba aproksimirati Tejlorovim redom do prvog stepena, u tački  $\mu_t$ . Neka je:

$$T(X_t) \sim T(\mu_t) + T'(\mu_t)(X_t - \mu_t), \quad (3.2.2)$$

gde je  $T'(\mu_t)$  prvi izvod od  $T(X_t)$  u tački  $\mu_t$ . Odatle je varijansa:

$$\text{Var}(T(X_t)) \sim (T'(\mu_t))^2 \text{Var}(X_t) = (T'(\mu_t))^2 cf(\mu_t). \quad (3.2.3)$$

Da bi tražena varijansa bila konstantna, stabilizaciona transformacija varijanse  $T(X_t)$  mora biti izabrana kao:

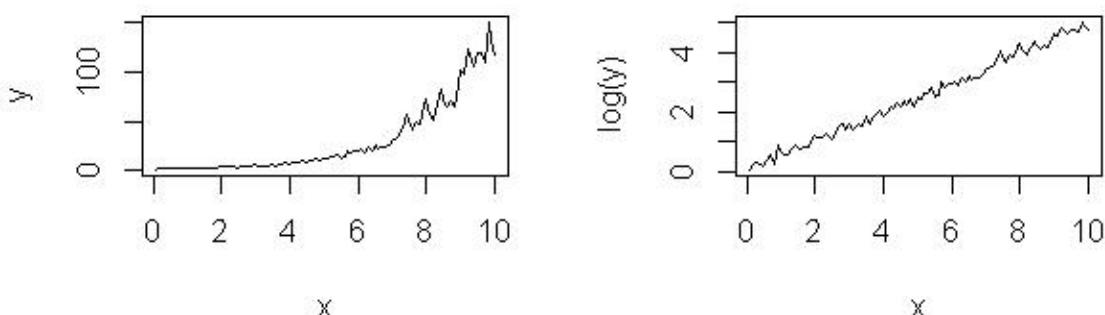
$$T'(\mu_t) = \frac{1}{\sqrt{f(\mu_t)}} \Rightarrow \int \frac{1}{\sqrt{f(\mu_t)}} d\mu_t. \quad (3.2.4)$$

Na primer, ukoliko je standardna devijacija proporcionalna nivou, tako da je  $\text{Var}(X_t) = c^2 \mu_t^2$  (gde je  $f(\mu_t) = \mu_t^2$ ), tada je:

$$T(X_t) = \int \frac{1}{\sqrt{\mu_t^2}} d\mu_t = \int \frac{1}{\mu_t} d\mu_t = \ln(\mu_t). \quad (3.2.5)$$

Dakle, logaritamska transformacija serije (osnova nije važna),  $\ln X_t$ , daje konstantnu varijansu.

Prostom simulacijom statističkim softverom R dobija se slika 3.2, na kojoj su grafik originalne serije i njene logaritamske transformacije. Jasno je da je varijansa stabilizovana i stacionarnost u odnosu na varijansu postignuta.



Slika 3.2: Logaritamska transformacija nestacionarne serije

### 3.3 ARIMA( $p, d, q$ ) procesi

**Definicija 3.3.1:** [10] Stohastički proces  $\{X_t, t \in T\}$  je *integriran proces reda d* ( $I(d)$ ) ako može biti transformisan u stacionaran stohastički proces diferenciranjem  $d$  puta.

Opštu klasu parametarskih modela u domenu vremena čine modeli *integriranih autoregresioneih pokretnih sredina* (ARIMA modeli), koji su pogodni za opisivanje stacionarnih, nestacionarnih, sezonskih i nesezonskih pojava. [7, 10, 2, 5, 30]

**Definicija 3.3.2:** [7] ARIMA( $p, d, q$ ) (engl. *autoregressive-integrated-moving-average*) proces dat je izrazom:

$$\phi_p(B)(1 - B)^d X_t = \theta_0 + \theta_q(B)\varepsilon_t, \quad (3.3.1)$$

pri čemu AR operator  $\phi_p(B) = 1 - \phi_1B - \phi_2B^2 - \dots - \phi_pB^p$ , reda  $p$  i invertibilni MA operator  $\theta_q = 1 - \theta_1B - \theta_2B^2 - \dots - \theta_qB^q$ , reda  $q$ , nemaju zajedničkih faktora, dok  $d$  označava nivo integriranosti vremenske serije. Za proces  $X_t$  kaže se da je integriran reda  $d$ , u oznaci  $I(d)$ . U praktičnim primenama, kod ekonomskih serija, obično je broj potrebnih razlika  $d$  jednak 0, 1 ili 2.

Za razliku od ARMA( $p, q$ ) modela, u izraz označen sa 3.3.1 uvedena je i konstanta  $\theta_0$ , koja ima različite uloge u slučaju kad je  $d = 0$  i u slučaju kad je  $d > 0$ . Ako je  $d = 0$ , proces je stacionaran, a konstanta je u relaciji sa sredinom procesa  $\theta_0 = \mu(1 - \phi_1 - \dots - \phi_p)$ , gde je  $\mu = E(X_t)$ . Međutim, u slučaju  $d \geq 1$ ,  $\theta_0$  se naziva član determinističkog trenda i često se izostavlja iz modela, osim ako nije zaista neophodan. Nadalje se pretpostavlja da koreni odgovarajućih karakterističnih jednačina  $\phi_p = 0$  i  $\theta_q = 0$  leže van jediničnog kruga. Drugim rečima, pretpostavlja se da su ispunjeni uslovi stacionarnosti i invertibilnosti.

Neki specijalni slučajevi ARIMA procesa: [5, 30, 7, 10]

1. ARIMA( $0, 1, 0$ ) - već pomenuti proces *slučajnog hoda*:

$$\Delta X_t = \varepsilon_t, \quad tj. \quad (1 - B)X_t = \varepsilon_t. \quad (3.3.2)$$

Ako se u ovaj izraz uvede konstanta  $\theta_0$ , 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* (engl. *random walk process with drift*).

2. ARIMA( $0, 1, 1$ ) - dat je izrazom:

$$\Delta X_t = \varepsilon_t - \theta_1\varepsilon_{t-1}, \quad tj. \quad (1 - B)X_t = (1 - \theta_1B)\varepsilon_t, \quad (3.3.3)$$

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. Parcijalni autokorelacioni koeficijent je samo na prvoj docnji 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:

$$\Delta^2 X_t = \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2}, \quad tj. \quad (1 - B)^2 X_t = (1 - \theta_1B - \theta_2B^2)\varepsilon_t \quad (3.3.4)$$

4. ARIMA( $1, 1, 1$ ) - dat je izrazom:

$$\Delta X_t - \phi_1\Delta X_{t-1} = \varepsilon_t - \theta_1\varepsilon_{t-1}, \quad tj. \quad (1 - \phi_1B)(1 - B)X_t = (1 - \theta_1B)\varepsilon_t \quad (3.3.5)$$

Sadašnja vrednost procesa  $X_t$  može se izraziti preko:

- prethodnih vrednosti  $X$ -eva i sadašnje i prethodnih vrednosti  $\varepsilon$ -a, direktnim korišćenjem diferencijalne jednačine:

Ako je  $\varphi(B) = \phi(B)(1 - B)^d = 1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_{p+d} B^{p+d}$ , a  $\theta_0 = 0$ , izraz 3.3.1 može se zapisati kao:

$$X_t = \varphi_1 X_{t-1} + \dots + \varphi_{p+d} X_{t-p-d} - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} + \varepsilon_t; \quad (3.3.6)$$

- beskonačne težinske sume sadašnje i prethodnih vrednosti  $\varepsilon_{t-j}$ :

$$X_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} = \psi(B) \varepsilon_t \quad (3.3.7)$$

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  $\varphi(B)X_t = \theta(B)\varepsilon_t \Rightarrow \varphi(B)\psi(B) = \theta(B)$ ;

- beskonačne težinske sume prethodnih vrednosti  $X$ -eva i sadašnje vrednosti  $\varepsilon_t$ :

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

$$X_t = \pi_1 X_{t-1} + \pi_2 X_{t-2} + \dots + \varepsilon_t \quad (3.3.8)$$

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

### 3.4 Izgradnja ARIMA modela

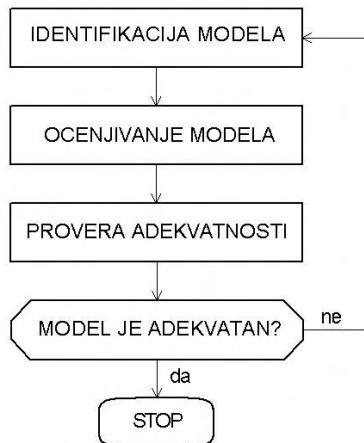
Opštu strategiju modeliranja ARIMA procesa koncipirali su *Box* i *Jenkins*, pa se prema njima ona naziva i *Box-Jenkinsova metodologija*. Kvalitetan model treba da prođe kroz tri etape ove metodologije: *identifikaciju modela*, njegovo *ocenjivanje* i *proveru* njegove *adekvatnosti* i da zadovolji osnovne principe koji karakterišu dobar model [7]:

1. *ekonomičnost* - opisati pojavu što jednostavnijim modelom, koji će istaći suštinsku karakteristiku 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. njegovo prilagođavanje podacima i apriornom znanju (ekonomskoj teoriji ili zdravom razumu);
4. *prihvatljivost podataka* - model ne sme da predviđa vrednosti koje ne zadovoljavaju neka definiciona ograničenja (npr. 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 (ako jedan od modela ima manju srednje kvadratnu grešku prognoze, a sve ostale karakteristike su jednake, tada se taj model kategorise kao prihvatljiviji za korišćenje);

6. *obuhvatnost* - model treba ne samo da objasni, odnosno opiše podatke, već i da objasni uspeh ili promašaj konkurentnog modela u objašnjenu istih podataka.

Osnovu *Box-Jenkins*-ove metodologije opšte strategije modeliranja čine tri etape izgradnje modela: [2, 7]

1. *identifikacija* - postupak korišćenja podataka vremenske serije u cilju izdvajanja uže klase ekonomičnih *ARIMA* modela, koji se uzimaju u razmatranje kao potencijalni generatori 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 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 3.3:** Dijagram toka *Box-Jenkinsovog iterativnog postupka*, [7]

### 3.5 Sezonski *ARIMA* modeli

Vremenske serije kod kojih se pojavljuju periodične fluktuacije u vremenskim intervalima do godinu dana, nazivaju se *sezonske vremenske serije*. Često se ovakve serije sreću u turizmu, proizvodnji, kretanjima zaliha itd., čija kretanja uslovjava vreme kao meteorološki fenomen ili institucionalna rešenja koja se primenjuju u privredi (npr. veći priliv turista u vreme praznika, povećani obim prodaje sladoleda u letnjim mesecima ili smanjenje proizvodnje usled štrajka ili kolektivnog odmora). 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 modelirati i sezonska kolebanja vremenske serije.

Najmanji vremenski period u kome se ponovi uočena pojava naziva se *period sezone* i označava se sa  $s$ . 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 [7].

Pored grafika sezonske vremenske serije, i koreogram sezonske vremenske serije može biti od pomoći pri utvrđivanju sezonskog karaktera serije. Lagano smanjivanje vrednosti autokorelacionih koeficijenata na sezonskim *lag*-ovima ( $s, 2s, 3s, \dots$ ) 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 3.5.1:** [7] Sezonski *ARIMA* model za seriju  $\{X_t, t \in T\}$  ima oblik:

$$\phi(B)(1 - B)^d(1 - B^s)^D X_t = \theta(B)\varepsilon_t, \quad (3.5.1)$$

gde su operatori nesezonskog i sezonskog diferenciranja primjenjeni  $d$ , odnosno  $D$  puta, respektivno.

Pri izgradnji sezonskog *ARIMA* modela primenjuje se iterativni postupak Boxa i Jenkinsa. Međutim, postoje izvesne poteškoće u primeni modela 3.5.1:

- Bar jedan od polinoma  $\phi(B)$  ili  $\theta(B)$  mora biti minimalno reda  $s$  da bi obuhvatio autokorelaciju na sezonskim *lag*-ovima. 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;
- Interpretacija parcijalne autokorelace 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* [7, 2].

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

$$(1 - B)^d \phi_p(B) X_t = \theta_q(B) \eta_t, \quad (3.5.2)$$

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

$$(1 - B^s)^D \Phi_P(B^s) \eta_t = \Theta_Q(B^s) \varepsilon_t, \quad (3.5.3)$$

gde su  $\Phi_P(B^s) = 1 - \Phi_1 B^s - \dots - \Phi_P B^{P_s}$  i  $\Theta_Q(B^s) = 1 - \Theta_1 B^s - \dots - \Theta_Q B^{Q_s}$  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 3.5.2 i 3.5.3 dobija se *Box Jenkinsov multiplikativni sezonski ARIMA model*:

$$(1 - B)^d(1 - B^s)^D \phi_p(B) \Phi_P(B^s) X_t = \theta_q(B) \Theta_Q(B^s) \varepsilon_t. \quad (3.5.4)$$

Uobičajeno je da se u modelu 3.5.4 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  $ARIMA(p, d, q) \times (P, D, Q)_s$ .

### 3.5.1 $ARIMA(0, 1, 1)(0, 1, 1)$ 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 vrednost 0 ili 1. [7, 2, 5, 10]

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

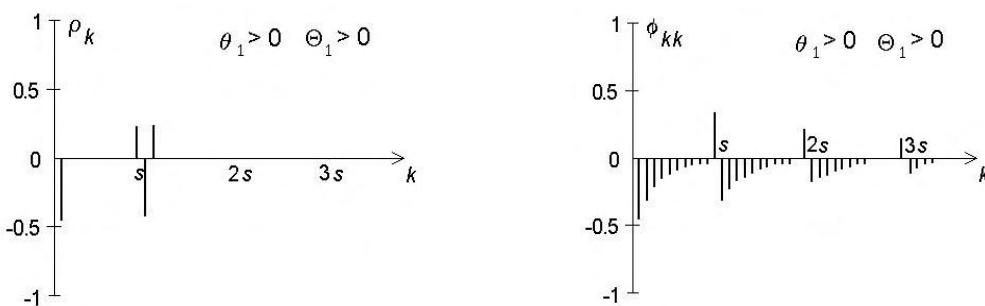
$$(1 - B)(1 - B^s)X_t = (1 - \theta_1 B)(1 - \Theta_1 B^s)\varepsilon_t, \quad (3.5.5)$$

koji je u literaturi poznat kao „vazduhoplovni model“ (engl. *airline model*). Invertibilnost ovog modela zahteva ispunjenje uslova  $|\theta_1| < 1$  i  $|\Theta_1| < 1$ .

Autokovarijaciona i autokorelaciona funkcija ovog modela su:

$$\gamma_k = \begin{cases} (1 + \theta_1^2)(1 + \Theta_1^2)\sigma^2 & , k = 0 \\ -\theta_1(1 + \Theta_1^2)\sigma^2 & , k = 1 \\ \theta_1\Theta_1\sigma^2 & , k \in \{s - 1, s + 1\} \\ -\Theta_1(1 + \theta_1^2)\sigma^2 & , k = s \\ 0 & , \text{inače} \end{cases} \quad \rho_k = \begin{cases} \frac{-\theta_1}{1 + \theta_1^2} & , k = 1 \\ \frac{\theta_1\Theta_1}{(1 + \theta_1^2)(1 + \Theta_1^2)} & , k \in \{s - 1, s + 1\} \\ \frac{-\Theta_1}{1 + \Theta_1^2} & , k = s \\ 0 & , \text{inače} \end{cases} \quad (3.5.6)$$

Za „vazduhoplovni“ model, na slici 3.4 prikazan je izgled običnog i parcijalnog koreograma za pozitivne vrednosti koeficijenata  $\theta_1$  i  $\Theta_1$ .



Slika 3.4: *Obična i parcijalna autokorelaciona funkcija „vazduhoplovnog“ modela*, [7]

## 4 DESEZONIRANJE VREMENSKIH SERIJA

### 4.1 Uvod

Makroekonomski statistike u kojima se podaci prikupljaju za razdoblje kraće od jedne godine, danas predstavljaju ključni alat za kreiranje ekonomske politike, analizu i modeliranje poslovnog ciklusa i prognoziranje. Međutim, ove statistike su često pod uticajem sezonskih fluktuacija i drugih učinaka kalendarata i radnih dana, koji mogu prikriti relevantna kratkoročna i dugoročna kretanja serije, te sprečiti jasno razumevanje ekonomske pojava.

Posledica toga je da su mnogi statistički podaci podložni sezonskom prilagođavanju, tj. *desezoniranju*. Glavna svrha desezoniranja je filtrirati uobičajene sezonske fluktuacije i tipične učinke kalendarata unutar kretanja analizirane vremenske serije. Uobičajene sezonske fluktuacije podrazumevaju ona kretanja koja se ponavljaju sličnim intenzitetom u istoj sezoni svake godine i za koja se na osnovu prošlih kretanja vremenske serije, pod normalnim okolnostima, može očekivati da će se ponavljati.

Fluktuacije koje nastaju zbog izrazito jakih ili slabih sezonskih uticaja (npr. zbog vremenskih uslova ili netičnih rasporeda praznika) i dalje će biti vidljive u desezoniranoj seriji prema stepenu u kojem premašuju ili se nalaze ispod normalnog sezonskog proseka. Jasno, ostali slučajni poremećaji i neobična kretanja koja se mogu razumeti u ekonomskim terminima (npr. posledice ekonomske politike ili štrajkovi) takođe će biti primetni.

Pored toga, desezoniranje obuhvata uklanjanje kalendarских efekata, jer uključuje uticaje koji su izvedeni iz razlika u broju radnih ili trgovackih dana (tzv. *working* ili *trading days*) ili određenih, statistički kvantifikovanih datuma (npr. državni praznici). Desezonirani podaci, zapravo, pomažu pri otkrivanju „novosti“ sadržanih u vremenskoj seriji, što je i krajnji cilj sezonskog prilagođavanja.

Statističke agencije, banke i druge institucije u svetu angažovane su svakodnevno na desezoniranju, te se mnogi resursi posvećuju filtriranju originalnih podataka. Zaista, desezoniranje je predmet stalne rasprave u mnogim aspektima, a mnoge metode i alati desezoniranja se još uvek razvijaju.

Rasprava se kontinuirano razvija, a još uvek se proučavaju novi smerovi i granice. Jedan od predmeta rasprave usredsređen je na prednosti i nedostatke desezoniranja, te na pridružene rizike „manipulacije“ podacima.

Sezonsko prilagođavanje (desezoniranje) se dokazalo kao koristan alat za ekonomsku analizu. Međutim, proizvođači podataka bi morali razmotriti sve prednosti i nedostatke desezoniranja i definisati jasnu strategiju proizvodnje pre početka ovog opsežnog procesa. Sezonsko prilagođavanje mora se sprovoditi samo kada postoji jasan statistički dokaz i ekonomska interpretacija sezonskih i kalendarских učinaka.

Treba biti svestan da sprovođenje bilo kog sezonskog i/ili kalendariskog prilagođavanja na serijama koje ne prikazuju dokaze takvih učinaka predstavlja neodgovarajući statistički tretman. Treba naglasiti da neke serije mogu biti karakterisane samo kalendarским učincima, bez sezonskih. U takvom slučaju biće prikladno samo *kalendarsko prilagođavanje*. Nadalje, druge serije mogu biti karakterisane samo sezonskim učincima, bez signifikantnih kalendarских učinaka. U tom slučaju treba primeniti samo filtriranje vezano za *sezonsko prilagođavanje*.

## 4.2 Prednosti, mere opreza, troškovi i rizici desezoniranja

Iz *Eurostat-a* (sedišta evropske statistike) i vrhunskih stručnjaka iz ove oblasti, dolaze jasna upozorenja o prednostima i rizicima pri desezoniranju, kao i preporuke o merama opreza, koje mogu pomoći u kvalitetnijem procesuiranju [15].

### Prednosti:

- pruža razumljive serije za analitičare otkrivajući „novosti“ sadržane u analiziranim vremenskim serijama;
- olakšava poređenje dugoročnih i kratkoročnih kretanja između sektora i zemalja;
- snabdeva korisnike potrebnim podacima za analizu poslovnog ciklusa, dekompoziciju trend-ciklusa i prepoznavanje tačaka obrta;
- primenjuje kontrolu kvaliteta i na strani *input-a* i na strani *output-a*, što omogućuje bolje poređenje s drugim serijama i metodama.

### Mere opreza:

- budući da sezonalnost nije precizno definisana, desezoniranje često zavisi od *a priori* prepostavki, sadržanih u osnovi odabranog modela i prepostavljenih procesa stvaranja podataka (subjektivnost sezonskog prilagođavanja). Sezonska i sezonski prilagođena komponenta mogu varirati od jednog do drugog softverskog programa, te zavisiti od odabranih opcija unutar softverskog programa;
- kvalitet desezoniranja u velikoj meri zavisi od kvaliteta originalnih (sirovih) podataka;
- niži nivo poređenja podataka između zemalja i statističkih područja ako se ne slede, odnosno nisu jasno definisana pravila i politike;
- mnoge ustanove zahtevaju proizvodnju velikog broja desezoniranih podataka i zbog toga je razumno da analitičar odluči o relativnoj važnosti serija korisnicima pre nego što doneše odluku o vremenu koje će posvetiti svakoj seriji;
- korisnici desezoniranih podataka moraju biti svesni toga da njihovo korišćenje u svrhe ekonometrijskog modeliranja treba pažljivo razmotriti.

### Troškovi i rizici desezoniranja:

- sezonsko prilagođavanje je vremenski intenzivno, te mu moraju biti posvećeni signifikantni računarski i ljudski resursi;
- zajednička i dobro definisana struktura IT-a za desezoniranje je bitan zahtev;
- sezonsko prilagođavanje koje je neadekvatno ili niskog kvaliteta može proizvesti rezultate koji mogu dovesti u zabluđu i povećati verovatnoću pogrešnih signala;
- prisutnost rezidualne sezonalnosti, kao i preterano izglađivanje predstavljaju konkretne rizike, koji mogu negativno uticati na interpretaciju sezonski prilagođenih podataka.

### 4.3 Istorija procesa desezoniranja

Analiza komponenti vremenskih serija pojavljuje se u radovima iz oblasti astronomije, meteorologije i ekonomije još u 17-om i 18-om veku, kao i u ranim istraživanjima koje je sprovodio *Buys-Ballot* (1847). Ti prvi radovi koncentrisani su prevashodno na uklanjanje „lažnih“ korelacija između dve promenljive. *Poynting* (1884) i *Hooker* (1901) pokušavali su da uklone trend iz cena pravljenjem proseka cena za nekoliko godina. *Spencer* (1904) i *Andersen* (1914) su uveli koncept polinoma višeg reda u eliminisanju trend komponente.

Ekspanziju razvijanja oblasti desezoniranja tokom dvadesetih i tridesetih godina prošlog veka uslovio je rad *W. M. Person-a* (1919), koji je osmislio metod za izolovanje četiri komponente iz vremenske serije:  $X_t = S_t \times T_t \times C_t \times R_t$  (prepostavljajući multiplikativnu vezu), gde je  $S_t$  sezonska komponenta,  $T_t$  trend,  $C_t$  ciklična, a  $R_t$  slučajna komponenta. Njegov metod koristi fiksirane sezonske faktore, iako je bilo jasno u literaturi tog doba da u pojedinim oblastima ideja o fiksnoj sezonalnosti nije bila validna. *Crum* je 1925. godine modifikovao *Person-ov* originalni metod tako da podržava i sezonske varijacije.

Prvu uopštenu metodologiju desezoniranja dao je *Macaulay* 1930. godine. Ovaj pristup se danas obično naziva „Klasična dekompozicija“, a postavio je temelje mnogim modernim pristupima, poput *X-11-ARIMA metode*.

Dva vrlo važna otkrića, koja se vezuju za rane pedesete godine prošlog veka, u velikoj meri doprinela razvojnom procesu metode desezoniranja. Prvo je uvođenje tehnike eksponencijalnog izravnjanja, koje je pojednostavilo dotadašnje dosadne proračune. Drugo je uvođenje računara, s obzirom na to da su računice koje su ranije izvođene danima, sada bile rešavane kroz nekoliko sekundi. Ovo je omogućilo razvoj raznih komplikovanih metoda, jer im je, testiranjem na velikom broju vremenskih serija, bilo lako utvrditi kvalitet.

**Census I metod** bio je predstavljen 1954. godine. Ovo prevazilazi *Macaley-ov* metod time što se prognoziranjem vrednosti pre početka serije (engl. *backcasting*) i po završetku serije (engl. *forecasting*) nadoknađuju vrednosti koje nedostaju. Ovo je bila osnovna verzija, a nju je *Shiskin* modifikovao 1955. godine u *Census II*, što je zapravo elektronska verzija ručne metode, koja se prethodno koristila.

**Census II metod** imao je nekoliko kritika, koje su dovele do stvaranja mnogo sofističiranije varijante od koje je nastao **X-11** 1965. godine, koji je najšire korišćeni metod desezoniranja do osamdesetih godina prošlog veka. *X-11* verzija imala je jedan vrlo važan napredak, tzv. *trading day prilagođavanje* pomoću različitih tehnika regresije. Pristup korišćen u *X-11* metodi rezultat je rada *Eisenpress-a* (1956), *Marris-a* (1960) i *Young-a* (1965). *X-11* takođe dozvoljava korisniku da odabere između aditivne i multiplikativne sezonalnosti, eliminaciju ekstremnih vrednosti iz dalje analize i sl. Njegovi nedostaci vezani su pre svega za korišćenje pokretnih proseka u postupku ocene trend-ciklusa, jer se u tim pokretnim prosecima koristi uvek isti skup pondera za sve vremenske serije, bez obzira na njihove individualne karakteristike.

Najpoznatiju izmenu metode *X-11* predložila je *Dagum*, predstavnica Statistike Kanade 1980. godine - tzv. **X-11-ARIMA**, koja se od svojih prethodnika razlikuje po izgradnji *ARIMA* modela, koji se potom koristi za prognozu prethodnih i budućih vrednosti posmatrane serije. Ovim prognozama se produžava analizirana serija, pa se postupak, koji je u osnovi metoda *X-11*, primenjuje na tako produženu vremensku seriju. Novije verzije (**X-11-ARIMA/88** i **X-11-ARIMA/2000**) proširile su metodologiju u smislu uvođenja dijagnostičkih poruka. To su *polu-parametarski modeli*.

Program **X-12-ARIMA** predstavio je 1997. godine američki Popisni biro (engl. *US Census Bureau*)<sup>6</sup>. Ima sve opcije kao i *X-11*, koji ocenjuje trend i sezonsku komponentu koristeći pokretnе proseke, ali nudi i mnoštvo novih mogućnosti uključujući: analizu trgovačkih dana (engl. *trading days*) i praznika, efekte autlajera (engl. *outlier*) i zamenu podataka koji nedostaju novim, a takođe je proširio spektar dijagnostike, kao što su istorija revizije (engl. *revision history*) i klizni rasponi (engl. *sliding spans*).

Pored metoda pokretnih proseka, razvijan je i čist *modelski pristup*, koji razlikuje:

- *Deterministički metod* - smatra trend i sezonalnost kao unapred određenu krivu ponašanja, a sve neizvesnosti su samo skretanja realnih vrednosti vremenske serije sa te krive. Ovaj model zasnovan je na regresionej analizi. Primeri determinističkih metoda su **DAINTIES** i **BV4**.
- *Stohastički metod* - pripisuje značajan efekat neizvesnosti i ima važnu ulogu u modeliranju. Razlikuju se dve glavne grupe globalnih stohastičkih modela: *ARIMA* model i Strukturni model.

Prema pristupu zasnovanom na *ARIMA* modelu (engl. *Auto Regressive Integrated Moving Average*) polazi se od modeliranja vremenskih serija, a onda se iz ovog procenjenog modela izvode modeli za komponente. U pristupu zasnovanom na struktturnom modelu polazi se direktno od ocenjivanja komponenata (*Engle (1978), Harvey i Todd (1983)*). Najprihvaćeniji strukturni modeli su **BAYSEA** (*Akaike i Ishiguro (1980)*), **DECOMP** (*Kitagawa (1985)*) i **STAMP** (*Koopman, Harvey, Doornik i Shepherd (1995)*).

Prva praktična realizacija *ARIMA* modeliranja potiče od Engleske banke (engl. *Bank of England*), iz osamdesetih godina 20-og veka. Dalji razvoj postigla je Španska banka (engl. *Bank of Spain*) pod kontrolom *Augustin Maravall-a*, što je kao rezultat dalo program **TRAMO/SEATS**<sup>7</sup> (*Gomez i Maravall (1997)*).

Program za desezoniranje **X-13-ARIMA-SEATS** (2006) je unapređena verzija *X-11* programa. Unapređenje uključuje svestran korisnički interfejs i raznovrsnost novih dijagnostika, koje pomažu korisniku da otkrije i otkloni nedostatke u sezonskom i kalendarskom prilagođavanju, dobijene na osnovu odabranih opcija u programu. Pored automatskog postupka modeliranja uvedenog u *X-11-ARIMA*, *X-13-ARIMA-SEATS* uključuje dodatnu proceduru modeliranja zasnovanu na *TRAMO* (engl. *"Time Series Regression with ARIMA Noise, Missing Observations and Outliers"*) metodi. *X-13-ARIMA-SEATS* takođe uključuje *SEATS* (engl. *"Signal Extraction in ARIMA Time Series"*) algoritam, koji dozvoljava korisnicima da primenjuju *regARIMA* model. Dok *X-11* metod izračunava dekompoziciju koristeći filtere pokretnih proseka za sezonom i trend, *SEATS* metod računa dekompoziciju koristeći filtere određene ocenjenim *regARIMA* modelom. Sa *X-13-ARIMA-SEATS* metodom omogućena je spektralna, revizijska i dijagnostika kliznih raspona (engl. *sliding spans*), koja je bila dostupna već duže u *X-11* tipu prilagođavanja, ali ne i u *SEATS* algoritmu.

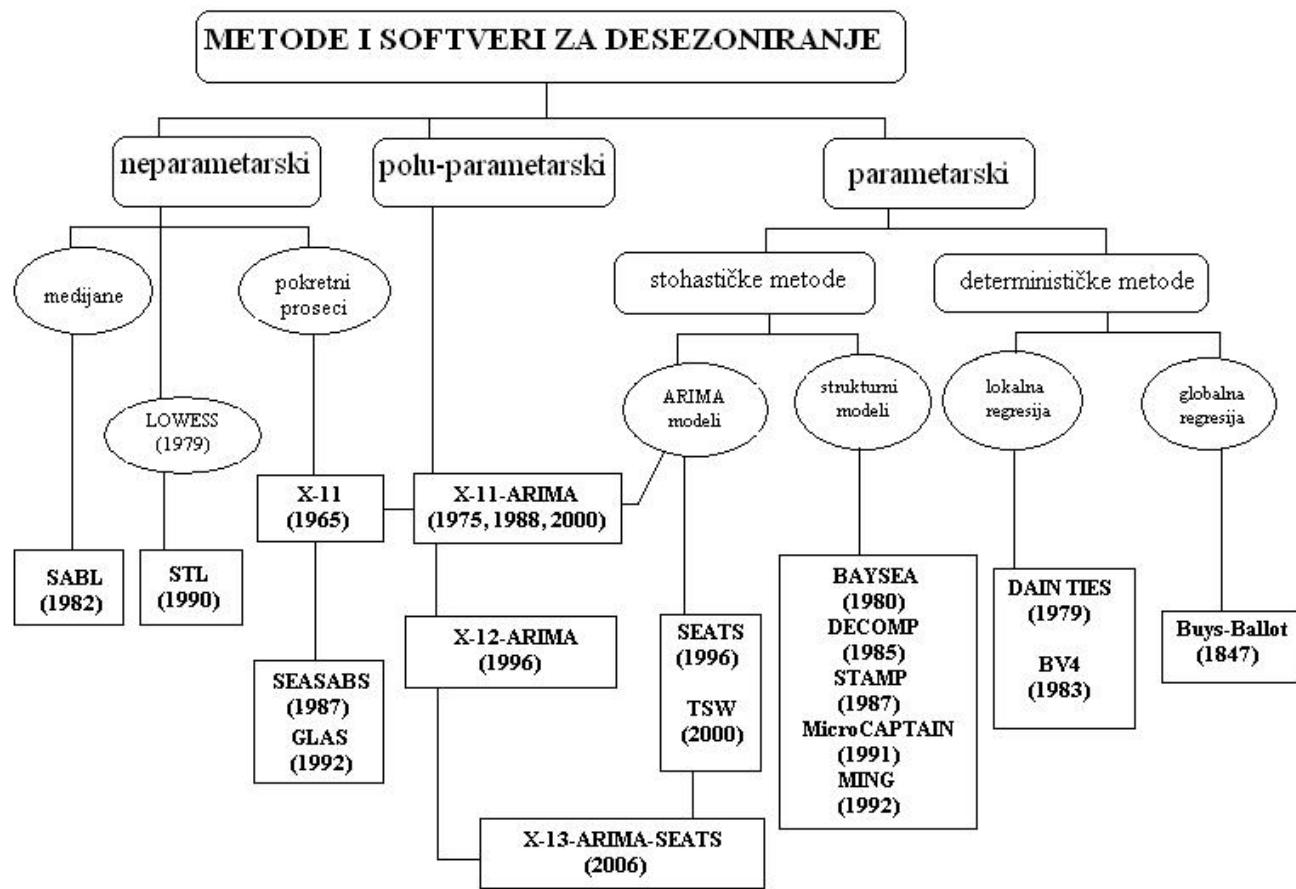
Pored parametarskih modela, razvijani su i neparametarski modeli, poput **SABL**-a (*Cleveland, Devlin i Terpenning (1982)*). To je metod koji je u principu sličan konstrukciji *X-11*. *SABL* je specijalno kreiran da podnese anomalije u podacima (autlajere). Može da radi sa podacima svih frekvencija (npr. mesečnim i kvartalnim), a izglađivanje funkcije kontroliše korisnik.

---

6. <http://www.census.gov>

7. <http://www.bde.es>

Razvoj metoda za desezoniranje najbolje je objašnjen na Slici 4.1.



Slika 4.1: Dijagram toka razvoja metoda za desezoniranje, [14]

Od devedesetih godina prošlog veka, *Eurostat* (Statistička kancelarija Evropske unije)<sup>8</sup> ima glavnu ulogu u promovisanju, razvoju i održavanju javno dostupnih softverskih rešenja za desezoniranje.

**Demetra** je alat koji je razvio *Eurostat*, a koji je bio vrlo pogodan za desezoniranje, jer je koristio *TRAMO/SEATS* i *X-12-ARIMA* metode. Komitet za monetarnu, finansijsku i statistiku platnog bilansa (CMFB) predstavio je smernice evropskog statističkog sistema (engl. *ESS Guidelines*), kao okvir za sezonsko prilagođavanje Glavnih evropskih ekonomskih indikatora (PEEI). Ove smernice predstavljaju korisne praktične preporuke za korišćenje i *TRAMO/SEATS* i *X-12-ARIMA* metode. Prvi korak je tzv. pred-podešavanje (engl. *Pre-adjustment*) u kom se uklanjaju deterministički efekti iz serije pomoću regresionih modela. Drugi deo je dekompozicija. Pristupi metoda *TRAMO/SEATS* i *X-12-ARIMA* u prvom koraku obrade vrlo su slični, ali se potpuno razlikuju u drugom koraku. Zato je vrlo teško porebiti rezultate, jer njihove dijagnostike počivaju na različitim aspektima, a izlazi imaju potpuno drugačije forme.

S obzirom na to da *Demetra* nije mogla da pokrije sve preporuke ESS smernica, Nacionalna banka Belgije ju je unapredila u aplikaciju **Demetra+**, od kad se potpuno prestaje sa korišćenjem prethodne verzije. Ova verzija je, međutim, počivala na korišćenju komponenti *FORTRAN* jezika i mogla se koristiti samo u *Windows* okruženju, te nije imala dugoročnu

8. <http://ec.europa.eu/eurostat>

perspektivu. Ovaj problem motivisao je *Eurostat* da preuzme inicijativu u kreiranju novog softvera za sezonsko prilagođavanje, koji će biti nezavistan od platforme.

Rezultat je plasiranje alata ***JDemetra+*** 2012. godine. Novi softver ima brojna poboljšanja u funkcionalnostima i mogućnostima koje nudi. Prvenstveno, ***JDemetra+*** omogućuje desezoniranje korišćenjem *X-13-ARIMA-SEATS* metode. Zatim, svi originalni *FORTRAN* kodovi prepisani su u *JAVA-i*, poštujući objektno orijentisani pristup. Ovakvo rešenje je od enormne važnosti, jer omogućuje dugoročno održavanje alata, integraciju biblioteke u IT okruženju mnogih institucija i ponovno korišćenje modula i algoritama za druge namene. [14, 22, 25, 26]

## 4.4 Postupak desezoniranja

Klasična analiza sezonskih vremenskih serija zasnovana je na razlaganju vremenske serije na trend, ciklus, sezonu i iregularnu komponentu. Prvobitno su trend i ciklična komponenta posmatrane razdvojeno. Međutim, u poslednje vreme one se posmatraju zajedno, tj. kao jedna komponenta. Zato će se u daljem tekstu govoriti o trendu, a podrazumevati trend-ciklus komponenta.

Pažnja se uglavnom usmerava na postupke otklanjanja sezonske komponente, tj. postupak sezonskog prilagođavanja, odnosno desezoniranja. Eliminacijom sezonske komponente prevashodno se želi omogućiti izučavanje suštinskih karakteristika vremenske serije, što ima svoju vrednost ne samo sa stanovišta istorijske ekonomske analize proteklih događaja, nego i zbog utvrđivanja tekućih privrednih uslova. One kreatorima ekonomske politike omogućuju preduzimanje pravovremenih mera u cilju stabilizovanja tekućih privrednih kretanja. Istovremeno, korisnici lakše interpretiraju desezonirane podatke, no serije kod kojih su ostale komponente „zamagljene“ prisustvom sezonske komponente.

### 4.4.1 Ciljevi pred-podešavanja serije (engl. *Pre-adjustment*)

Većina metoda za sezonsko prilagođavanje i softverskih programa procenjuje sezonsku komponentu upotrebom linearnih postupaka i filtera: modeli *ARIMA*, pokretni proseci, regresiona analiza itd. Ti linearni alati su optimalni pod preciznim pretpostavkama, ali imaju i neke nedostatke [15]:

- nisu otporni, tj. osjetljivi su na prisutnost netipičnih vrednosti (autlajera);
- nisu postojani, tj. osjetljivi su na bilo koju pogrešnu specifikaciju osnovnog modela.

Glavni cilj pred-podešavanja serije je osigurati pouzdanu procenu sezonske komponente. To se radi uglavnom pronalaženjem i ispravljanjem serija za podatke i/ili komponente, koje se ponekad nazivaju „nelinearnosti“, a mogu sprečavati procenu sezonalnosti.

Autlajeri su jasan primer podataka koji mogu snažno uticati na kvalitet procene sezonalnosti. *TRAMO-SEATS*, *X-12-ARIMA* i *X-13-ARIMA-SEATS*<sup>9</sup> imaju ugrađene postupke pronalaženja i ispravljanja serija za različite vrste autlajera: impulsne autlajere (engl. *additive outliers*), prolazne promene (engl. *transitory changes*) i pomake nivoa (engl. *level shifts*).

---

9. Zbog glomaznog zapisa, u daljem tekstu *X-12-ARIMA* imaće označku *X-12*, a *X-13-ARIMA-SEATS* označku *X-13*.

Ekonomski vremenski serije se obično beleže svakog meseca (ili kvartala), ali meseci (ili kvartali) nisu jednaki. Oni uglavnom nemaju ni jednaku dužinu, ni isti sastav prema broju dana. Te specifičnosti, strogo povezane sa kalendarom, mogu uticati na sirove podatke. Na primer, jedna subota više u mesecu može objasniti povećanje prometa u maloprodaji. *TRAMO-SEATS*, *X-12* i *X-13* imaju ugrađene postupke prepoznavanja i ispravljanja serija za ove kalendarске efekte (npr. trgovачke dane i efekat Uskrsa). Takođe, oni mogu uzeti u obzir i nacionalne specifičnosti (npr. verski i građanski praznici) pomoću specifičnih regresora i težinskih koeficijenata.

Treba primetiti da je deo kalendarског ућinka sezonski (dužina većine meseci ponavlja se svake godine, pravoslavni Uskrs pada češće u aprilu, nego u martu i maju, itd.), pa se ispravljanje izraženo kalendarском komponentom treba odnositi samo na nesezonski deo efekta, dok sezonski deo uticaja kalendaru treba pripisati sezonskoj komponenti. Takođe je važno primetiti da analitičar ima vrlo malo dvoumljenja oko budućnosti kalendaru, jer se ponavlja u razdoblju od 2800 godina. Zbog toga će dobra procena efekta trgovачkih dana voditi boljom prognozi originalnih podataka i stabilnijim sezonskim procenama.

Ostali efekti, kao što su temperatura, školski praznici, prenošenje praznika itd., mogu objasniti kratkoročno ponašanje serije, no pod normalnim uslovima seriju ne treba korigovati za te učinke. Treba odlučivati od slučaja do slučaja. Uticaj ovih učinaka može se analizirati za posebne potrebe. Uglavnom je učinak temperature u velikom procentu sezonski, a budućnost nesezonskog dela je vrlo neizvesna. Pod ovakvim uslovima, uzimanje u obzir takvog efekta u modeliranju serije može rezultirati velikim revidiranjem najnovijih vrednosti desezonirane serije.

Većina statističkih alata upotrebljenih u postupcima desezoniranja zavisi, barem u jednom koraku prilagođavanja, od stacionarnosti serije. Stacionarnost prema sredini može se obično postići diferenciranjem. Stacionarnost prema varijansi može zahtevati transformaciju serije. *TRAMO-SEATS*, *X-12* i *X-13* pružaju korisniku automatski test za log-transformaciju. Rezultat testa će takođe voditi ka odabiru sheme dekompozicije.

Pristup kome treba težiti je sprovođenje detaljnog pred-podešavanja, tj. detaljni prethodni tretman serije. Prihvatljiva i održiva opcija, naročito ako prva zahteva angažovanje mnogo resursa za prilagođavanje velikih količina podataka ili kako bi se savladale specifične karakteristike podataka, je sprovođenje potpuno automatskog pred-podešavanja. Praksa koju treba izbegavati je da uopšte nema prethodnog tretmana serije.

#### 4.4.2 Grafička analiza serije

Prva grafička analiza serije pruža analitičaru neke korisne informacije o tome kako sprovesti desezoniranje i odabrati parametre, i otkriva moguće probleme u podacima. Tu analizu moguće je sprovesti koristeći se osnovnim grafikonima ili pomoću prvog prolaza programa za desezoniranje koristeći zadate (engl. *default*) parametre. Tada se mogu prikupiti sledeće informacije: [15, 19, 21]

- dužina serije (serije kraće od tri godine ne mogu se prilagođavati pomoću preporučenih programskih paketa za desezoniranje);
- prisutnost neobičnih vrednosti (npr. nule ili autlajeri) ili problemi s podacima;
- struktura serije: prisutnost trenda, sezonske komponente, volatilnosti itd.;
- prisutnost mogućih prekida u sezonskom ponašanju;

- shema dekompozicije (aditivna, multiplikativna).

Složeniji grafikoni, kao što je spektralni grafikon ili autokoreogram, mogu pružiti informacije o prisutnosti sezonske komponente i/ili učinku trgovačkih dana. Ta prva analiza može biti vremenski zahtevna kada se mora analizirati vrlo veliki broj serija, no relevantna je barem za najvažnije serije.

Naravno, uvek se prvo preporučuje detaljna grafička analiza, temeljena na osnovnim grafikonima, autokoreogramima i spektralnim prikazima, barem za najvažnije serije koje treba desezonirati. Tu analizu treba dopuniti prvim eksploratornim prolazom programa za desezoniranje na celom skupu. Ako se radi o velikom broju serija i nedostatku vremena, alternativna opcija je samo prva grafička analiza bar na najvažnijim serijama, (a kad god je moguće, na svim) po mogućnosti koristeći se prvim prolazom programa za desezoniranje. Treba izbegavati nesprovodenje eksploratorne analize važnih serija. U tom slučaju kvalitet desezoniranja u velikoj se meri oslanja na tačnost i relevantnost automatskih postupaka i testova za prepoznavanje i ispravljanje problema u podacima.

#### 4.4.3 Kalendarsko prilagođavanje

Ljudi su podložni prirodnom ritmu. Jasno je koliko faktori kao što su trajanje dana, promena godišnjih doba i promenljivi vremenski uslovi utiču na aktivnost. Zimi, na primer, troši se više energije zbog grejanja. Na nas utiču društveni ritmovi, kao što su, na primer, praznici. Ukoliko se meri rezultat ljudske aktivnosti svakog meseca ili kvartala, jasno je da u merenjima važan trag ostavljaju ovakvi prirodni i društveni ritmovi. [15, 23, 24, 26, 29]

Struktura i sastav kalendara mogu uticati na ekonomске aktivnosti na različite načine. Ti učinci, ako za njih nije učinjena odgovarajuća ispravka, mogu uzrokovati pogrešnu specifikaciju modela *ARIMA* i ugroziti ukupan kvalitet desezoniranja. Efekti kalendara obično obuhvataju:

- različit broj radnih dana u određenom mesecu ili razdoblju - meseci generalno ne sadrže ceo broj nedelja; svaki dan u nedelji u jednom mesecu ponovi se najmanje četiri puta, a može se javiti i peti put; kako mnoge aktivnosti zavise od nedeljnog ritma, ovaj uticaj je očigledan; primer je recimo proizvodnja televizora (ako mesec ima 22 radna dana, verovatno će biti više televizora, nego ako mesec ima 21 radni dan)
- sastav radnih dana - neke institucije imaju neradni ponedeljak, a posluju vikendom;
- učinak prestupne godine - ako je prestupna godina, februar ima 29 dana, što automatski ima uticaj na kretanje serije;
- pokretne praznike, kao što su neki nacionalni praznici, Uskrs, Ramazan, itd. - praznici generalno utiču na vrednost podataka u vremenskim serijama, jer se većina računa kao slobodni dan, a to se odražava na, na primer, proizvodnju ili prodaju.

Važno je primetiti da je deo kalendarskih efekata sezonski, tako da se uklanja standardnim filterima za sezonsko prilagođavanje. Kalendarsko prilagođavanje unutar pred-podešavanja desezoniranja mora obrađivati samo nesezoniski deo gore spomenutih efekata.

Preporučuje se sprovođenje kalendarskog prilagođavanja na svim serijama koje prikazuju značajne i uverljive kalendarske učinke, u okviru postojanog statističkog pristupa, raspoloživog u najnovijim alatima za sezonsko prilagođavanje. U svakom slučaju, treba izbegavati nikakvo prilagođavanje kad postoji dokaz kalendarskih efekata, ili kalendarsko prilagođavanje svih serija bez ikakve provere važnosti i uverljivosti učinaka [15].

## 1. Metode prilagođavanja za trgovačke/radne dane (engl. *trading/working days*)

Cilj prilagođavanja za trgovačke/radne dane je dobijanje sezonski prilagođene serije čije su vrednosti nezavisne od dužine i sastava dana u nedelji (broj ponedeljaka, utoraka itd./broj radnih dana i dana u vikendu) za određeni mesec, odnosno kvartal.

Treba napomenuti da su dužina i sastav dana u nedelji za mesec/kvartal delimično sezonski: mart uvek ima 31 dan i u proseku sadrži više ponedeljaka nego februar. Budući da je sezonski deo već zahvaćen filterima za desezoniranje, ne treba ga uklanjati u postupku kalendarskog prilagođavanja. Stoga, efekat radnih/trgovačkih dana - u užem smislu - mora biti pridružen nesezonskom delu učinka.

Efekat prestupne godine (engl. *leap year*) je nesezonski deo učinka dužine meseca, odnosno kvartala. Nesezonski deo sastava dana u nedelji za mesec/kvartal može se proceniti pomoću odstupanja broja radnih ili trgovačkih dana od njihovog dugoročnog mesečnog, odnosno kvartalnog proseka. Prilično tačne procene ovih dugoročnih proseka mogu se izračunati u kalendaru čija je dužina deljiva sa 28 godina.

Prilagođavanje za radne ili trgovačke dane preporučuje se za seriju u kojoj se takvi efekti mogu pronaći, a pri tom su uverljivi i statistički bitni. Pod pretpostavkom da se upotrebljava prihvaćena metoda prilagođavanja (obično ona koja se temelji na procjenjenom modelu), to bi svakako trebalo da redukuje ili potpuno ukloni efekte u poslednjim godinama. Metoda ne bi trebalo da rezultira čestim velikim revizijama kada novi podaci postanu dostupni. Ako se to dogodi, postoji indikacija da procene dobijene korišćenom metodom nisu pouzdane.

Najčešće se preporučuje pristup pomoću modela regARIMA, sa svim prethodnim testovima za broj regresora, dužinu i sastav meseca, s proverom uverljivosti efekata (predznak i veličina procenjenih koefficijenata) itd. Naravno, prilagođavanje treba sprovoditi u onim vremenskim serijama u kojima postoje statistički dokazi i ekonomsko objašnjenje za postojanje kalendarskih učinaka. Treba izbegavati proporcionalno prilagođavanje (efekat trgovačkih dana procenjuje se brojanjem njihove proporcije u mesecu, odnosno kvartalu) ili nikakvo prilagođavanje u slučajevima kada postoji dokaz učinaka trgovačkih dana u prilagođenoj seriji, jer bi to verovatno uticalo na modeliranje serije pomoću *ARIMA* modela.

## 2. Prilagođavanje za pokretne praznike (engl. *moving holidays*)

Cilj prilagođavanja za pokretne praznike je dobijanje sezonski prilagođene serije, čije su vrednosti u pojedinačnim tačkama nezavisne od pojedinih efekata kalendara, koji prate kompleksan obrazac tokom godina. Katolički i pravoslavni Uskrs, npr., mogu imati različite efekte na susedne mesece ili kvartale, te tako uzrokovati probleme pri interpretaciji podataka u mesecima i kvartalima na koje utiču. Zbog toga je, kada se takvi efekti pronađu u vremenskoj seriji i nisu zanemarljivi, preporučljivo pokušati ih ispraviti, tj. izračunati i ukloniti procene njihovih efekata.

Ti efekti mogu biti delimično sezonski: katolički Uskrs, na primer, češće pada u aprilu nego u martu. Budući da je sezonski deo već izolovan pomoću filtera za desezoniranje, ne treba ga uklanjati pri kalendarskom prilagođavanju. Obično se nesezonski deo efekta pokretnog praznika može proceniti uklanjanjem njegovog dugoročnog mesečnog/kvartalnog proseka iz regresora.

I ovde se preporučuje pristup regARIMA, s prethodnim testovima za efekte Uskrsa i drugih pokretnih praznika. Dužine efekta pokretnih praznika definišu se na osnovu

prethodnih testova i provere uverljivosti učinaka. Treba izbegavati netestiranje, odnosno ispravak za gore pomenute efekte, ako dijagnostički dokazi ne potvrđuju potrebu za tim.

### 3. Nacionalni i kalendar EU-a/evrozone

Nacionalni i kalendar EU-a/evrozone mogu se koristiti za kalendarsko prilagođavanje, kako bi se uzele u obzir nacionalne pojedinosti. Neki programi za desezoniranje, naročito *TRAMO-SEATS*, *X-12* i *X-13*, dopuštaju integraciju tih kalendara ili putem regresora ili kao ponuđenu funkcionalnost za otkrivanje i ispravljanje kalendarskih efekata.

Kalendar EU-a/evrozone, izgrađen iz nacionalnih kalendara, tj. uprosećivanjem nacionalnih brojeva radnih i trgovačkih dana, korišćenjem odgovarajućih težinskih koeficijenata, može se smatrati alternativom u slučajevima direktnog sezonskog prilagođavanja originalnih agregata EU-a. Mogu se koristiti dodatni nacionalni kalendari u gore spomenutim alatima za desezoniranje. Nije lak zadatak napraviti i održavati nacionalne kalendare i evropski kalendar, a njihov učinak u velikoj meri zavisi od redovnog i korektnog održavanja.

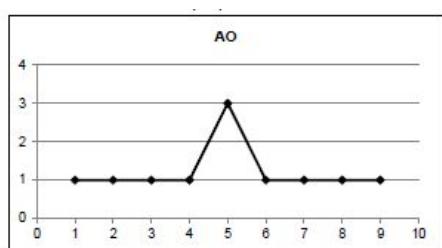
Zemlje članice se ohrabruju da sastave, održavaju i ažuriraju svoje nacionalne kalendare, ili kao minimalnu alternativu da pruže istorijski spisak državnih praznika, jer neraspoloživost takvim informacijama može da ugrozi sposobnost Eurostata da ostvari cilj smernica za kalendarsko prilagođavanje.

Upotreba nacionalnih kalendara preporučuje se na nivou zemlje ili za evropske aggregate kada je odabran indirektni pristup. Upotreba kalendara EU-a/evrozone preporučuje se kada je odabran direktni pristup za desezoniranje evropskih agregata, naročito ako serije prilagođene pomoću nacionalnog kalendara nisu raspoložive, nepotpune su ili su nedovoljno statistički kvalitetne. Upotrebљene kalendarske informacije treba da su na raspolaganju javnosti (bar na zahtev).

Alternativa je upotreba zadatih (engl *default*) kalendara (definisanih u sklopu odabranog alata za sezonsko prilagođavanje), dopunjenih listom nacionalnih praznika, a svakako se ne preporučuje upotreba zadatih (engl. *default*) kalendara, bez ikakvog pozivanja na nacionalne i evropske državne praznike, te nesprovođenje kalendarskog ispravljanja uprkos dijagnostičkim dokazima kalendarskih efekata.

#### 4.4.4 Pronalaženje i ispravljanje ekstremnih vrednosti (engl. *Outliers*)

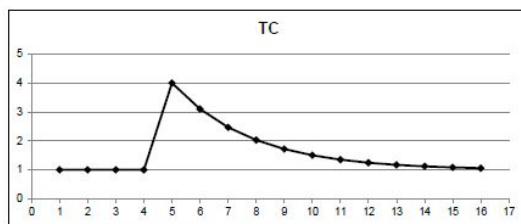
Ekstremne vrednosti su vrednosti serije koje se naročito ističu. Postoji nekoliko tipova autlajera, a neki od najčešćih su: [23, 21, 15]



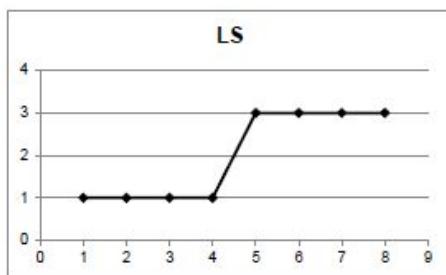
Slika 4.2: Aditivni autlajer

- **impulsni ili aditivni autlajeri** (engl. *additive outliers - AO*) - ekstremne vrednosti u izolovanim tačkama serije; na grafikonu se vidi kao vrh, jer vremenska serija odjednom ima značajno povećanje ili smanjenje, a zatim se odmah vraća na svoju prethodnu vrednost, tako da se uticaj oseti samo na jednom podatku; javlja se u slučaju kolektivnog odmora, štrajka itd.

- **prolazne promene** (engl. *transitory changes - TC*) - niz autlajera s prolaznim efektima na nivo serije; vremenska serija se zbog nekih 'smetnji' iznenada promeni (poraste ili opadne), a zatim se postepeno vraća u stanje pre poremećaja; razlozi su slični kao kod *AO*, samo što je ovde potrebno više vremena da njihov uticaj nestane.



**Slika 4.3:** *Transitory changes autlajer*



**Slika 4.4:** *Level shift autolajer*

- **pomaci nivoa** (engl. *level shifts - LS*) - niz autolajera sa konstantnim i trajnim efektom na nivo serije; primetan porast ili pad kojim se promeni nivo serije i ta promena ostaje; nastaje usled promena u definicijama, klasifikaciji ili metodologiji.

Metode sezonskog prilagođavanja, koje se uobičajeno temelje na linearnim modelima, verovatno će biti pod znatnim uticajem takvih autlajera (ako su prisutni), te ih zbog toga treba pronaći i zameniti pre procenjivanja sezonske i kalendarske komponente, kako bi se izbegla njihova iskrivljena ili pristrasna procena. Međutim, autlajeri treba da ostanu vidljivi u sezonski prilagođenim podacima (osim ako ne predstavljaju greške u podacima), jer pružaju informacije o nekim posebnim događajima (kao što su, na primer, štrajkovi).

S autlajerima nije lako postupati, posebno na kraju serije, kada je teško razlikovati ih od tačaka obrata (engl. *turning point*). *TRAMO-SEATS*, *X-12* i *X-13* sadrže automatski postupak za pronađenje autlajera i ispravljanje njihovog efekta. Skraćivanje vremenskog raspona i promena kritične vrednosti statističkih testova mogu pomoći u dobrom modeliranju autlajera. Postupak koji treba izbegavati je potpuno ignorisanje njihovog postojanja [15].

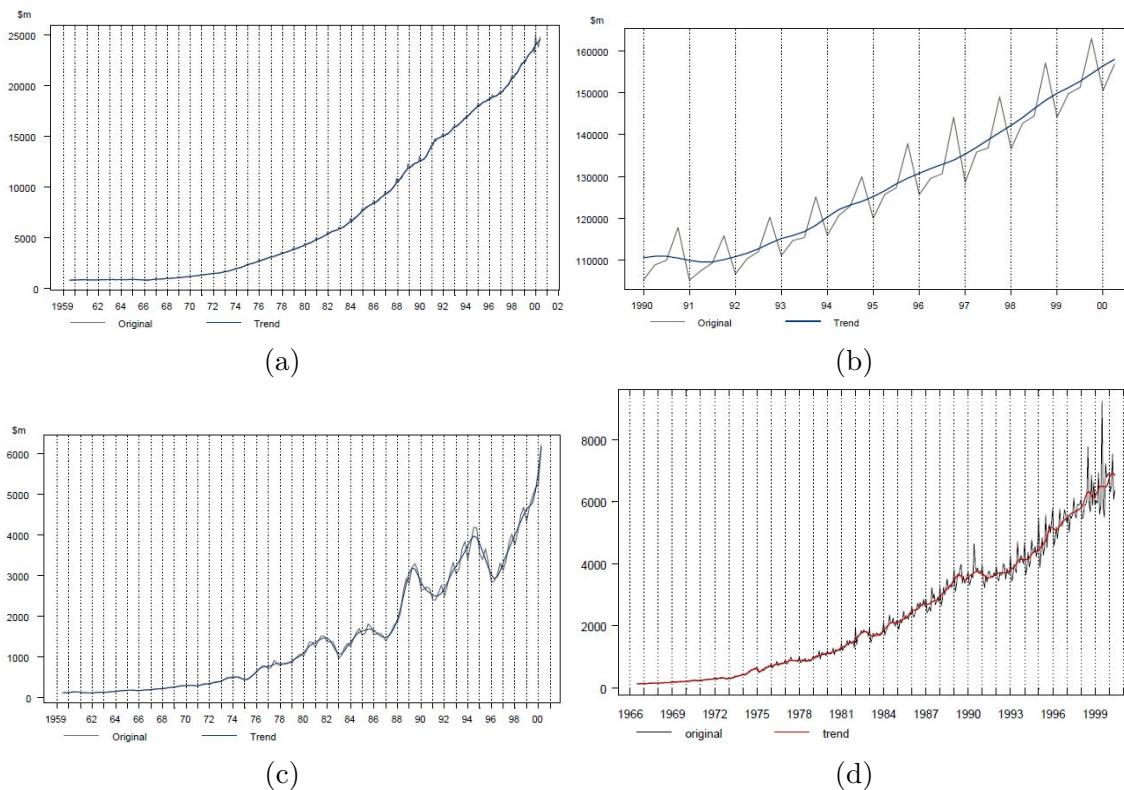
#### 4.4.5 Shema dekompozicije

Shema kompozicije podrazumeva način na koji se različite komponente - u osnovi trend-ciklus, sezonska i iregularna - kombinuju, kako bi oblikovale originalnu seriju [23, 28, 29, 15, 16].

### 1. Trend-ciklus komponenta (engl. *trend-cycle component*)

Definiše se kao dugoročni prosek kretanja ili pravac razvoja vremenske serije. Može biti pozitivan ili negativan, u zavisnosti od toga da li serija pokazuje tendenciju u rastu ili opadanju. Ako nema uzlazne ili silazne putanje, serija je stacionarna u sredini. Metode za analizu vremenske serije često procenjuju trend i neke ciklične fluktuacije zajedno. Stoga se jednakom koriste izrazi „trend-ciklus“ i „trend“. Ciklične varijacije koje spadaju u trend-ciklus komponentu su one koje se javljaju u periodu dužem od godinu dana (nekad i do nekoliko desetina godina). Ako se posmatra previše kratka vremenska serija, ciklične fluktuacije se poistovećuju sa trendom, pa je stoga u definiciji trenda izuzetno važan pojam „dugoročno kretanje“. Neki primjeri vremenskih serija sa izraženom trend komponentom dati su na slici 4.5 [16]:

- a) Poslednjih 40 godina nivo serije se postepeno i konstantno povećava. Originalni podaci i trend su na istom grafiku i gotovo su identični, što ukazuje na vrlo malo sezonskih i uticaja iregularnih podataka. Sveukupno dugoročno kretanje serije ima tendenciju rasta.
- b) Vrlo sezonska serija, koja takođe ima rastući trend tokom vremena, a koji pokazuje opšte dugoročno kretanje serije u odsustvu sezonskih i iregularnih faktora. U ovom slučaju trend je generalno rastući, sa blagim padom početkom devedesetih godina.
- c) Trend je doživeo nekoliko tačaka preokreta, jer prati ciklus podataka u sredini. Vizuelna sugestija, međutim, je da je u proteklih 40 godina dugoročno zabeležen postepeni porast.
- d) I ovde je uočljiva tendencija rasta. Međutim, nepravilnosti u seriji, koje se mogu interpretirati kao slučajna kretanja oko trenda, takođe rastu tokom vremena. Ovo sugerise da je količina nepravilnog kretanja u vezi sa nivoom trenda, pa je najbolje koristiti multiplikativni model.

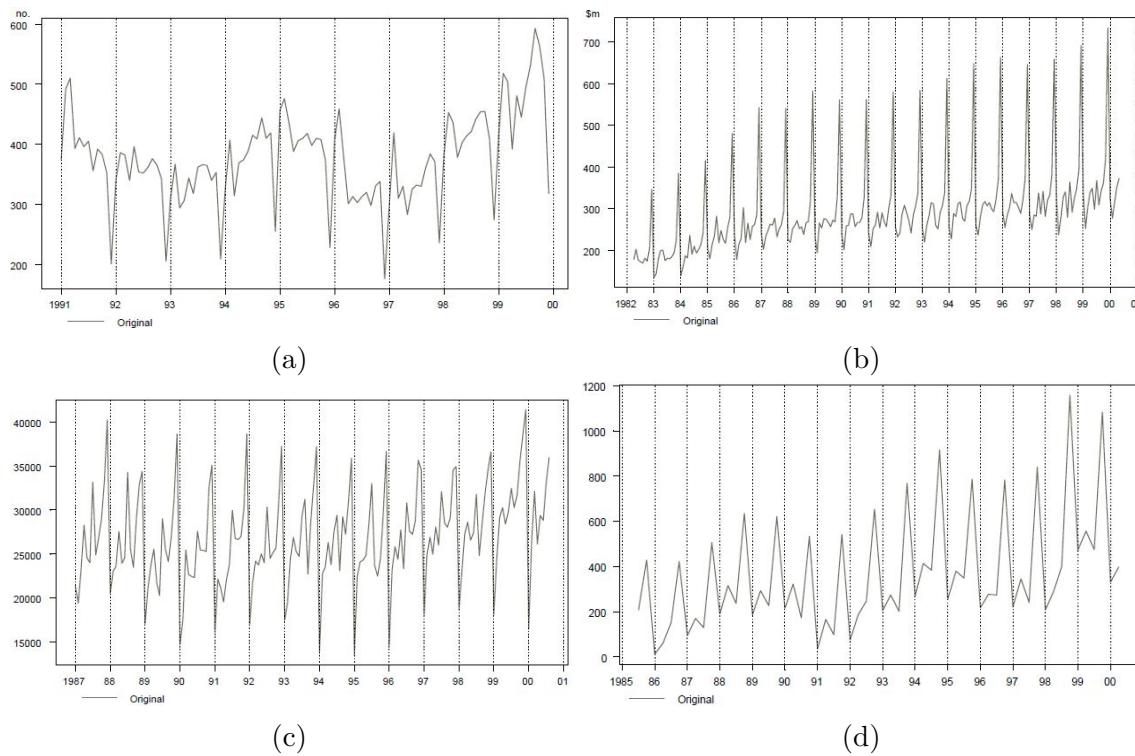


**Slika 4.5:** Primeri vremenskih serija sa izraženim trendom, [16]

## 2. Sezonska komponenta (engl. seasonal component)

Ova komponenta obuhvata sezonske i kalendarske efekte, što uključuje prirodne i društvene uticaje, kao što su promena godišnjih doba, broj radnih dana i uticaj praznika. Sezonska komponenta ima period od godinu dana i relativno je stabilna u odnosu na vreme, pravac i opseg. Može se identifikovati kao pravilno raspoređeni usponi i padovi (svaka dvanaesta opservacija za mesečne i svaka četvrta opservacija za kvartalne serije), koji su doslednog pravca i približno istog opsega svake godine. Kretanja sezonskih faktora odražavaju se na unutarnjopravno istočno promene u trajanju vremenske serije. Sledeći primeri predstavljaju neke tipične vremenske serije koje pokazuju jaku sezonsku komponentu. Interpretacija slike 4.6 [16]:

- a) U decembru svake godine zabeležen je sezonski pad u poslovnim programima, verovatno zbog sezone godišnjih odmora. Na grafiku se ovaj efekat odražava kroz oštре redovne „uvale“ u seriji.
- b) Očigledan je veliki sezonski rast u maloprodajnoj kupovini u decembru, zbog Božićne kupovine i mnogo manji sezonski porast tokom sredine godine. Oštре šiljci na grafiku su potvrda ove pojave.
- c) Sezonski obrazac odražava ciklus potrošnje. Uprkos nekim nepravilnostima u seriji, sezonski porasti i dalje mogu biti vidljivi u decembru, a sezonski padovi svakog januara.
- d) Ovo je kvartalna serija koja pokazuje maloprodajni profit. Godišnji obrazac je često lakše utvrditi kod kvartalnih serija, jer se ponavlja svake četvrte opservacije. U ovom slučaju postoji sezonski porast profita u četvrtom kvartalu svake godine.



**Slika 4.6:** Primeri vremenskih serija sa izraženom sezonom, [16]

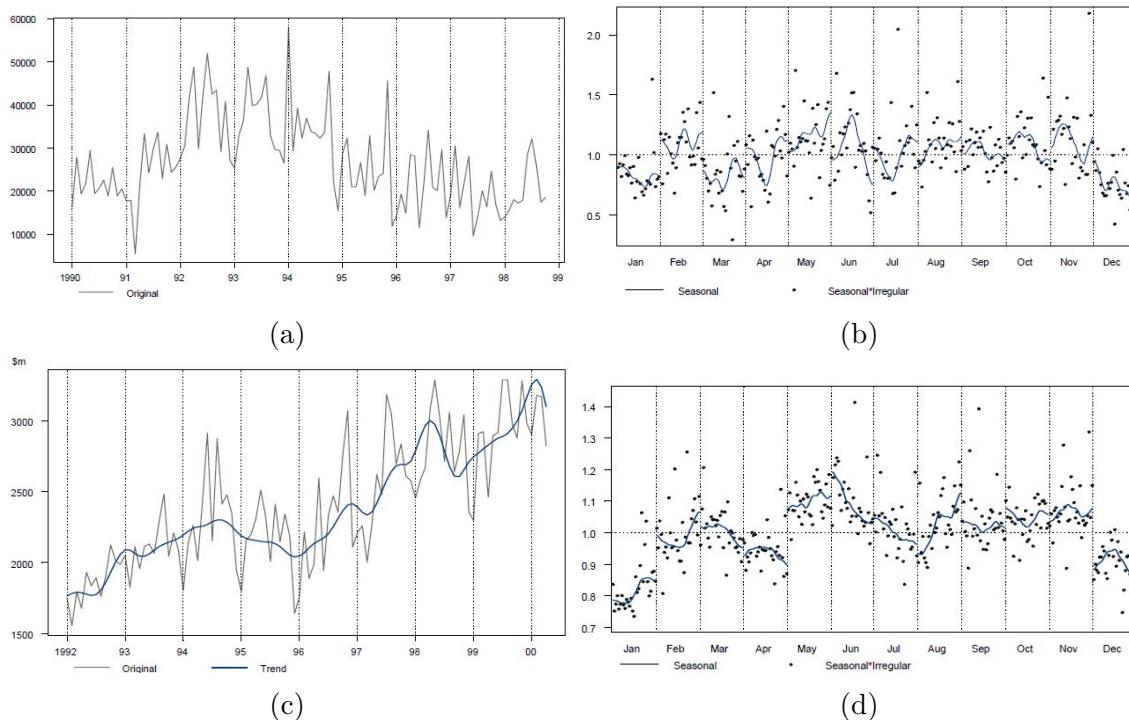
### 3. Iregularna komponenta (engl. irregular component)

Sastoje se od fluktuacija koje ne pripadaju ni trend-ciklusu, ni sezoni. Ovo su ostaci,  $AO$  i  $TC$  autolajeri i poneke prolazne fluktuacije. Iregularna komponenta je bitna u proceni validnosti modela. Ako se izabere dobar model, ostaci predstavljaju slučajnu grešku. U proseku, ova komponenta treba da ima vrednost 0 ako je model aditivan ili 1 ako je model množilični. Vrednosti u različitim vremenskim tačkama moraju biti nezavisne jedne od drugih, što znači da se ne mogu predvideti na osnovu prethodnih. Većina serija sadrži izvestan stepen volatilnosti, izazivajući oscilacije originalnih i desezoniranih vrednosti oko opšteg nivoa trenda. Međutim, u slučaju kad je stepen irregularnosti neobično veliki, vrednosti mogu mnogo odstupati od trenda, što daje ekstremnu vrednost. Nepravilnosti u seriji odražavaju se kao kratkoročne oscilacije trenda, koja ne prate ni sistematsko kretanje, ni ponavljanje. Kod visoko nepravilnih

serija, reziduali mogu da dominiraju i da uzrokuju težu identifikaciju sezone i trenda. Tada se koriste SI grafikoni (engl. *seasonal-irregular chart*), na kojima se vide sezonski i sezonsko-irregularni faktori serije za određeni mesec (poredak je po godini).

Na slici 4.7 prikazane su irregularne komponente i njihovi SI grafici za dve vremenske serije. Interpretacija grafika je sledeća [16]:

- Ovo je primer vrlo nepravilne serije. S obzirom na to da je vrlo teško utvrditi da li su kratkoročna kretanja uzrokovana sezonskim ili irregularnim uticajima, mora se koristiti i SI grafikon ove serije.
- Tačke u SI grafikonu predstavljaju sezonsko-irregularne, a puna linija sezonske faktore izračunate za seriju. Ovde se može uočiti široko variranje sezonsko-irregularnih oko sezonskog faktora, što ukazuje na dominaciju irregularne komponente u seriji.
- Prikazana je originalna mesečna serija izgradnje i građevinskih dozvola u Australiji u određenom periodu, kao i njena trend linija. Veliki uticaj imala su odobrenja za izgradnju Melburn kazina (jun, 1994), Sidnej kazina (aprili, 1995) i Olimpijskog stadiona (novembar, 1996), što je izazvalo neuobičajeno visoke ekstreme u seriji i učinilo je neregularnom.
- SI grafik ove serije posebno naglašava ekstreme u junu, aprilu i novembru, što potvrđuje uticaj nastale situacije. Ovakve ekstreme je teško otkriti i na SI graficima. Najbolje su uočljivi na zajedničkom grafiku za desezoniranu seriju i trend komponentu. Razlog za to je što ova serija nije naročito sezonska, pa su originalni i desezonirani podaci vrlo slični i irregularni. Otuda se desezonirana serija, koja sadrži i irregularnu i trend komponentu, značajno razlikuje od trenda. Kod nestabilnih serija je takođe teško otkriti ekstreme na SI grafiku, jer postoji mnogo varijabilnosti u sezonskoj i irregularnoj komponenti.



**Slika 4.7:** Primeri vremenskih serija sa izraženom irregularnom komponentom, [16]

#### 4.4.6 Odabir modela

Odluka o odabiru modela je subjektivna. Model je samo aproksimacija stvarnih podataka, tako da se kasnije po potrebi može menjati. Radije se koriste jednostavniji modeli, jer se lakše interpretiraju, a i mogu se koristiti za kratke vremenske serije. Takođe, može se na bilo kom delu vremenske serije izabrati posebno model, a onda oni zajedno formiraju model cele serije. Ponekad se mogu odabratи nekoliko modela (ako je, na primer, kod svih podjednako dobro uklapanje s podacima), pri čemu se različiti modeli koriste u različite svrhe.

Odabir modela zapravo se odnosi na kriterijume za odabir adekvatnog modela za pred-prilagođavanje i sezonsko prilagođavanje, *log* nasuprot *ne-log*, tj. *none* specifikaciji modela, red diferenciranja za sezonski i nesezonski deo, upotrebu aditivnih i multiplikativnih procena modela, analizu dekompozicije na osnovu odabranog modela itd. Relevantnost ove stavke je bitno različita ako je reč o metodama baziranim na modelu ili neparametarskim metodama. Nepoznati parametri se ocenjuju metodom najmanjih kvadrata ili metodom maksimalne verodostojnosti.

Pri izboru modela treba obratiti pažnju na sledeće [15, 16, 19, 20, 21] :

##### 1. Autokoreaciona funkcija (engl. *autocorrelation function*)

Autokorelacija, tj. koeficijent autokorelacije definisan je, kako je u poglavljiju 1.2 naznačeno, sa  $\rho_k = \frac{\text{Cov}(X_t, X_{t-k})}{\text{Var}(X_t)}$ , za  $k \geq 0$ . Ukoliko se autokorelacija posmatra kao funkcija promenljive, ova funkcija se zove autokoreaciona funkcija (*ACF*). Iz *ACF* se može lako zaključiti koliko je neka vrednost povezana sa prethodnim vrednostima. Na primer, ako  $\rho_1$  meri kako je podatak vremenske serije u vezi sa prethodnim vrednostima, onda  $\rho_{12}$  meri kako se podatak povezuje sa dvanaestim podatkom ispred njega. Autokorelacija se ocenjuje na osnovu uzorka autokorelaciјe, tj. uzorka koeficijenta autokorelaciјe, koji se izračunava po formuli:

$$r_k = \frac{\frac{1}{N-k} \sum_{t=1}^{N-k} (X_t - \bar{X})(X_{t+k} - \bar{X})}{\frac{1}{N} \sum_{t=1}^N (X_t - \bar{X})^2}, \quad \text{za } k \geq 0. \quad (4.4.1)$$

Za veliko  $N$ , faktor  $\frac{1/(N-k)}{1/N}$  je blizu 1. Ako se uzorak autokoreacione funkcije posmatra kao funkcija promenljive, zove se uzoračka autokoreaciona funkcija (uzoračka *ACF*). Grafik funkcije *ACF*, tj. uzoračke *ACF* zove se autokoreogram. Ako je vremenska serija stacionarna, za  $k > 0$   $r_k$  je približno normalno raspodeljena sa očekivanjem  $-\frac{1}{N}$  i varijansom  $\frac{1}{N}$ . Iz toga sledi da oko 95% uzoračke autokorelaciјe leži na intervalu  $(-\frac{1}{N} - \frac{2}{\sqrt{N}}, -\frac{1}{N} + \frac{2}{\sqrt{N}})$ . Približno se uzima  $(-\frac{2}{\sqrt{N}}, \frac{2}{\sqrt{N}})$ .

Uzoračka *ACF* pomaže pri izboru odgovarajućih modela. Posebno se kod mesečnih vremenskih serija mora voditi računa o  $r_1, r_2, r_{12}, r_{24}$ , odnosno o  $r_1, r_2, r_4, r_8$  kod kvartalnih vremenskih serija. Ako se ovi koeficijenti nalaze izvan gornjeg intervala ili blizu granica, onda treba pokušati sa popravkom modela.

##### 2. Pogodnost ***ARIMA*** modela

Da bi korisnik bio siguran u adekvatnost odabranog modela, treba da prođe kroz dodatne faze provere:

- utvrđivanje da li interpretacija modela odgovara smislu podataka;
- provjeravanje statističke značajnosti parametara modela;

- testiranje statističke značajnosti dodatnih parametara (npr. ako smo se odlučili za  $ARMA(p,q)$  model, testira se statistička značajnost dodatnih parametara u  $ARMA(p+1,q)$  i  $ARMA(p,q+1)$  modelu);
- provera uravnoteženosti  $ARIMA$  modela ( $p + P + d + D = q + Q$ );
- provera da li se mogu eliminisati sve ostale informacije (sezonske, kalendarske), a da ostaci budu slučajni, sa približno normalnom raspodelom;
- dodatni kriterijumi - ako je  $N$  broj podataka korišćenih za procenu modela (tj. broj podataka u diferenciranoj vremenskoj seriji),  $n_p$  broj ocenjenih parametara,  $L_n$  prilagođena funkcija maksimalne verodostojnosti, definiše se:

$$\begin{aligned} & - AIC = -2L_N + 2n_p \\ & - AICC = -2L_N + 2n_p \left(1 - \frac{n_p+1}{N}\right)^{-1} \\ & - BIC = -2L_N + n_p \ln \ln N \end{aligned}$$

Najbolji je model kod kojeg je vrednost kriterijuma najmanja.

### 3. Spektralna analiza

Spektar, odnosno spektralna gustina je nenegativna funkcija, simetrična na intervalu  $[-\pi, \pi]$ , pa je dovoljno posmatrati je na intervalu  $[0, \pi]$ . Dovoljan uslov za postojanje spektralne gustine je da je suma apsolutnih vrednosti autokovarijacije konačna. U praksi, to važi za većinu vremenskih serija. Ako se uzmu teorijske autokovarijacije, dobija se teoretski spektar. On se ocenjuje pomoću dostupnih podataka. Vrhovi, koji se pojavljuju u spektru, sugerisu da su za vremensku seriju karakteristične periodične oscilacije, koje odgovaraju dатој frekvenci.

Spektar se lako računa za svaku komponentu vremenske serije. Srazmeran je varijansi komponente. Kad su spektralni vrhovi uski, komponenta je stabilna, a ako su široki, ona je nestabilna (sadrži više slučajnih varijabilnosti).

Spektar trend-ciklus komponente ima spektralni vrh (engl. *peak*) na frekvenciji 0.

Spektar sezonske komponente ima spektralne vrhove zavisno od vrste serije:

- Ako se posmatraju mesečne vremenske serije, čiji je period 12, prva sezonska frekvanca se dobija na sledeći način:

$$\frac{2\pi}{perioda} = \frac{2\pi}{12} = \frac{\pi}{6} = 30^\circ. \quad (4.4.2)$$

Stoga, na intervalu  $[0, \pi]$  postoji šest frekvenci:  $\frac{\pi}{6}, \frac{\pi}{3}, \frac{\pi}{2}, \frac{2\pi}{3}, \frac{5\pi}{6}, \pi$ .

- Ako se posmatraju kvartalne vremenske serije, čiji je period 4, prva sezonska frekvanca dobija se na sledeći način:

$$\frac{2\pi}{perioda} = \frac{2\pi}{4} = \frac{\pi}{2} = 90^\circ. \quad (4.4.3)$$

Stoga, na intervalu  $[0, \pi]$  postoje dve frekvence:  $\frac{\pi}{2}, \pi$ .

Spektar iregularne komponente (belog šuma) je konstantan.

Spektar desezonirane vremenske serije jednak je zbiru spektara trend-ciklus i iregularne komponente.

Spektar vremenske serije jednak je zbiru spektara desezonirane vremenske serije i sezonske komponente. Ako je za vremensku seriju značajna sezona, spektar originalne vremenske serije ima pikove na jednoj ili više sezonskih frekvenci.

Ako spektar nema vrhove na sezonskim frekvencama, onda sezonske varijacije verovatno nisu dovoljno značajne, pa se za takve vremenske serije obično biraju nesezonski modeli. U desezoniranim vremenskim serijama ne sme biti sezone, što znači da se ne smeju pojaviti vrhovi na sezonskim frekvencama. Ako se pri indirektnom desezoniranju zapazi da sezona nije dobro odstranjena, mora se upotrebiti direktno desezoniranje. Ako testiranje ne pokaže značajne efekte kalendarata, ali se zapazi vrh na prvoj odgovarajućoj frekvenci u spektru ostataka ili neki sličan razlog, treba uključiti efekat kalendarata.

#### 4. Efekat sezone

Može se desiti da efekti sezone za vremenske serije nisu značajni, ali u slučaju da jesu, na grafu se vide značajne oscilacije, sezonska komponenta u izabranom modelu pokazuje smislene sezonske fluktuacije, spektar ima pik na sezonskim frekvencijama...

#### 5. Efekat kalendarata

Može se desiti da je efekat kalendarata za vremensku seriju značajan. U spektru se pri izabranom modelu vidi pik na frekvencama značajnim za uticaj kalendarata. Efekat kalendarata se uključuje u model samo ako je to statistički opravdano.

#### 6. Slične i srodne vremenske serije

Slične i srodne vremenske serije obično imaju slične modele.

#### 7. Više odgovarajućih modela

Ako se pronađe više odgovarajućih modela, mora se odlučiti za jedan od njih.

#### 8. Problem krajnje tačke (engl. *end-point problem*)

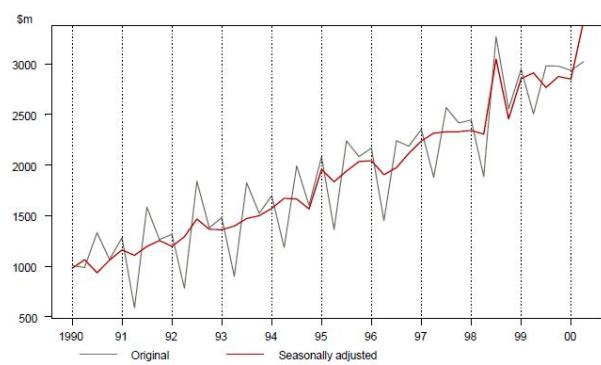
Pri desezoniranju se krajnje tačke vremenskih serija obično ocenjuju korišćenjem asimetričnih filtera, pa su zato procene manje pouzdane i može se desiti da se budući podaci znatno promene. Takav je slučaj, na primer, trend kod koga krajnja tačka signalizira rast, a sledeći podatak pokaže pad vrednosti serije.

Najbolja opcija je automatski odabir unutar velikog broja modela prema opcijama alata, nakon provere prikladnosti modela upotreboom standardnih statističkih testova (normalnost, heteroskedastičnost, serijska korelacija itd.) i spektralna dijagnostika. Nakon toga, radi se ručni odabir modela za važne i problematične serije. Prihvatljiva alternativa je upotreba potpuno automatskog postupka, a treba izbegavati odabir koji se temelji na ograničenom broju predefinisanih modela, koji nisu testirani na prikladnost skupu serija koje se prilagođavaju.

Najčešće korišćeni modeli dekompozicije su [16]:

- **Aditivni model**

Ovaj model dekompozicije pretpostavlja da se komponente serije ponašaju nezavisno jedna od druge. Na primer, povećanje trend-ciklusa neće izazvati povećanje sezonske komponente. Koristi se ako su iregularna i sezonska komponenta nezavisne od kretanja trenda, odnosno ako su efekti sezone isti iz godine u godinu. Tipičan aditivni model prikazan je na slici 4.8 ([16]). Trend serije varira, a amplitudne podešene serije približno su iste.



Slika 4.8: Aditivni model dekompozicije

Matematički, originalna serija  $X_t$  predstavlja zbir svojih komponenti (sezonske  $S_t$ , trend-ciklus  $T_t$  i iregularne  $I_t$ ) u trenutku  $t$ :

$$X_t = T_t + S_t + I_t, \quad za \quad \forall t \quad (4.4.4)$$

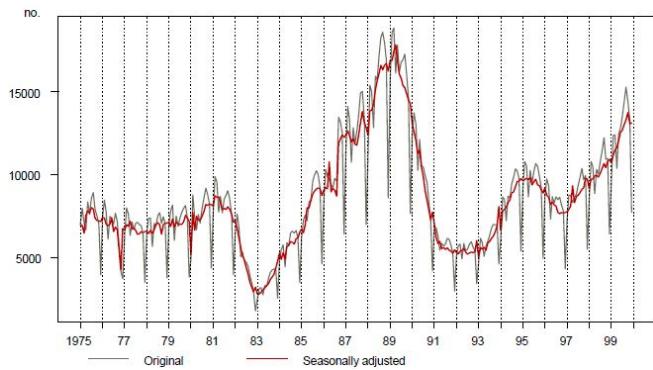
Desezonirana vemenska serija dobija se oduzimajem sezonske komponente od originalnih podataka ili sumiranjem trend-ciklus i iregularne komponente:

$$SA_t = X_t - S_t = T_t + I_t, \quad za \quad \forall t. \quad (4.4.5)$$

Proseci sezonskih fluktuacija su kod aditivnog modela oko nule tokom cele godine, tj  $\sum S_{t+j} = 0$ , gde je  $j = 1, \dots, n$ , a  $n = 12$  za mesečne, odnosno  $n = 4$  za kvartalne serije.

### • Multiplikativni model

Kod većine ekonomskih vremenskih serija veličine sezonske komponente variraju proporcionalno nivou serije, te je ovaj model dekompozicije najčešće korišćen. Tipičan multiplikativni model prikazan je na slici 4.9 ([16]). Ako trend serije raste, amplitude sezonskih „uvala i šiljaka“ takođe rastu.



**Slika 4.9:** Multiplikativni model dekompozicije

Matematički, originalna serija multiplikativnog modela,  $X_t$ , predstavlja proizvod svojih komponenti (sezonske  $S_t$ , trend-ciklus  $T_t$  i iregularne  $I_t$ ) u trenutku  $t$ :

$$X_t = T_t \cdot S_t \cdot I_t, \quad za \quad \forall t \quad (4.4.6)$$

U ovom modelu, desezonirana vemenska serija dobija se deljenjem originalnih podataka sezonskom komponentom ili množenjem trend-ciklus i iregularne komponente:

$$SA_t = X_t / S_t = T_t \cdot I_t, \quad za \quad \forall t. \quad (4.4.7)$$

Kod multiplikativnog modela zahteva se ograničenje  $\sum S_{t+j} = n$ , gde je  $j = 1, \dots, n$ , a  $n = 12$  za mesečne, odnosno  $n = 4$  za kvartalne serije. Ovo obezbeđuje da sezonski faktori ne obuhvataju dugoročne cikluse u vremenskoj seriji. Na primer, ako sezonski faktori ne bi bili centrirani oko jedinice tako da  $\sum S_{t+j} \neq n$ , gde je  $n$  broj perioda u toku godine, onda bi oni obuhvatili neke od godišnjih promena u osnovnom nivou serije. Dakle, ovakvim ograničenjem sezonskog faktora osigurava se da se oni u proseku kreću oko jedinice i samim tim ne mogu da doprinesu nivou serije tokom godine. To znači da kada je jedan period sezonski veći (sezonski faktor  $> 1$ ) od ostalih perioda, mora da postoji bar jedan drugi period koji nadoknađuje to tako što je sezonski nizak (sezonski faktor  $< 1$ ), da bi zbir tokom 12 meseci (4 kvartala) ostao i dalje  $n$ .

Multiplikativni model se logaritamskom transformacijom može dovesti u aditivnu formu:

$$\log X = \log T_t + \log S_t + \log I_t \quad (4.4.8)$$

To je **log-aditivni** model. Jasno je da multiplikativni i log aditivni model ne mogu biti korišćeni kada originalni podaci sadrže nulu kao vrednost.

U principu se model multiplikativne dekompozicije može izvesti iz aditivnog modela ako se pretpostavi da je svaka aditivna sezonska komponenta direktno proporcionalna nivou trenda, npr. za konstantu  $k_S$ , i da aditivne iregularne komponente imaju standardnu devijaciju direktno proporcionalnu kombinovanoj sumi trenda i sezonske komponente. Matematički zapis ova dva uslova je sledeći:

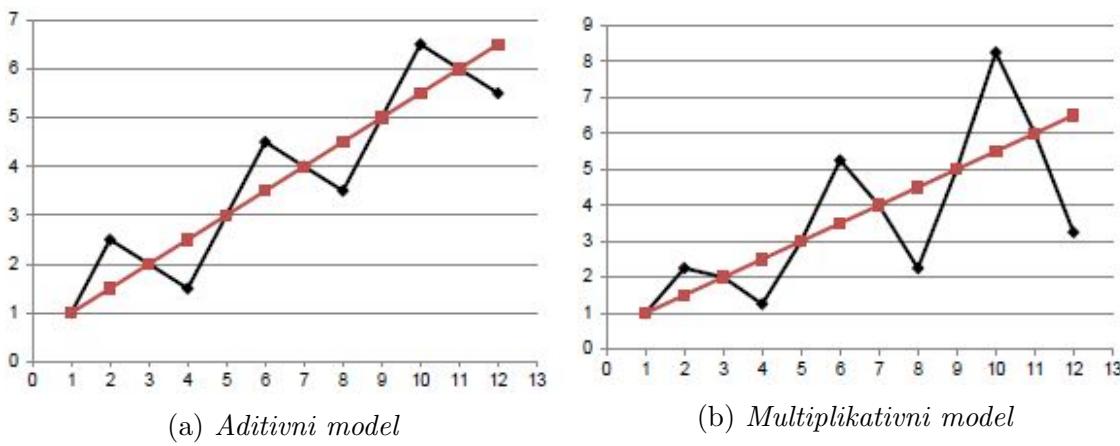
$$S_t^{(a)} = k_S T_t^{(a)} \quad i \quad I_t^{(a)} = K_I(t)(T_t^{(a)} + S_t^{(a)}), \quad (4.4.9)$$

gde je  $K_I(t), t = 0, 1, \dots, n$  niz slučajnih veličina sa očekivanjem nula i konstantnom varijansom, a oznake  $(a)$  i  $(m)$  razlikuju parametre aditivnog i multiplikativnog modela, respektivno. Korišćenjem ove dve pretpostavke, aditivni model se može zapisati kao:

$$\begin{aligned} X_t &= T_t^{(a)} + S_t^{(a)} + I_t^{(a)} \\ &= T_t^{(a)} + S_t^{(a)} + K_I(t) (T_t^{(a)} + S_t^{(a)}) \\ &= (T_t^{(a)} + S_t^{(a)}) (1 + K_I(t)) \\ &= (T_t^{(a)} + k_S T_t^{(a)}) (1 + K_I(t)) \\ &= T_t^{(a)} (1 + k_S) (1 + K_I(t)) \\ &= T_t^{(m)} \cdot S_t^{(m)} \cdot I_t^{(m)} \end{aligned} \quad (4.4.10)$$

Ovo je multiplikativni model sa trendom  $T_t^{(m)} = T_t^{(a)}$ , sezonskim faktorom  $S_t^{(m)} = 1 + k_S$  i iregularnim faktorom  $I_t^{(m)} = 1 + K_I(t)$ .

Jasna razlika između serija za koje treba koristiti aditivni, odnosno multiplikativni model dekompozicije prikazan je na sledećoj slici. U oba slučaja, prikazane su kvartalne serije originalnih podataka (grafik označen crnom bojom) i trend komponenta (grafik označen crvenom bojom).

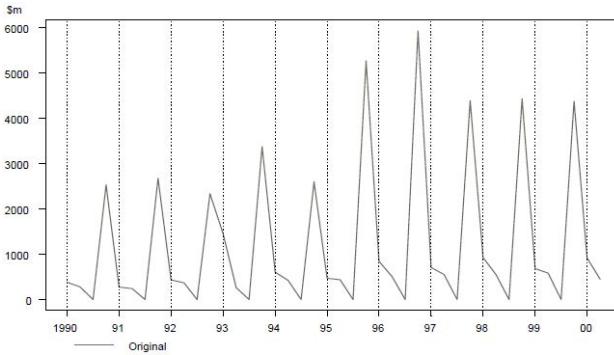


Slika 4.10: Razlika između aditivnog i multiplikativnog modela

### • Pseudo-aditivni model

Ovaj model dekompozicije kombinuje elemente i aditivnog i multiplikativnog modela. Pretpostavlja se da su sezona i iregularni elementi zavisni od nivoa trenda, ali da su oni međusobno nezavisni. Uklanjanje multiplikativne veze između sezonske i iregularne komponente je naročito korisno ako serija sadrži jednu ili više nula vrednosti. Razlog za to je što aditivna veza omogućava da bilo sezonska ili iregularna komponenta apsorbuju

nula vrednosti bez narušavanja trenda. Primer kvartalne serije koja zahteva pseudo-aditivnu dekompoziciju data je na slici 4.11, na kojoj je prikazana proizvodnja žitarica i to samo tokom određenih meseci, dok je proizvodnja praktično nula za četvrtinu svake godine.



Slika 4.11: Pseudoaditivni model dekompozicije, [16]

Algebarski, originalni podaci mogu se izraziti kao:

$$X_t = T_t + T_t \cdot (S_t - 1) + T_t \cdot (I_t - 1) = T_t \cdot (S_t + I_t - 1) \quad (4.4.11)$$

Desezonirana vemenska serija definisana je sa:

$$SA_t = X_t - T_t \cdot (S_t - 1) = T_t \cdot I_t \quad (4.4.12)$$

I kod pseudo-aditivnog modela, kao i kod multiplikativnog, sezonski i iregularni faktori su centrirani oko jedinice. Zato se mora oduzeti 1 od  $S_t$  i  $I_t$  da bi se osiguralo da su izrazi  $T_t(S_t - 1)$  i  $T_t(I_t - 1)$  centrirani oko nule. Ovi se izrazi mogu interpretirati kao aditivna sezonska i aditivna iregularna komponenta i, s obzirom na to da su centrirane oko nule, originalna serija  $X_t$  biće centrirana oko vrednosti trenda,  $T_t$ .

Aditivni i log-aditivni modeli deo su i TRAMO-SEATS, X-12 i X-13 metode za dekompoziciju, a multiplikativni je deo samo X-12 i X-13 metode. TRAMO-SEATS, X-12 i X-13 pružaju korisniku automatski test za log-transformaciju. Rezultat ovog testa takođe će sugerisati odabir sheme dekompozicije. Za seriju s vrednostima nula ili negativnim, automatski je odabrana aditivna dekompozicija pomoću postupka desezoniranja, bez obzira na stvarnu shemu dekompozicije. Za serije koje sadrže trendove i u sredini i u varijansi (prisutnost heteroskedastičnosti), log-aditivna dekompozicija je najprikladnija. Ako je prisutan samo trend u sredini, očigledno se koristi multiplikativna dekompozicija [15].

Preporučuje se automatski odabir dekompozicije upotrebom odgovarajućeg kriterijuma (npr. informacioni kriterijum) nakon grafičkog ispitivanja serije. Za problematične serije treba koristiti ručni odabir. Upotrebu fiksne sheme dekompozicije (npr. multiplikativne za pozitivne serije, aditivne za nepozitivne serije) treba potpuno izbegavati [15].

U narednoj tabeli dat je sažeti prikaz sva tri modela dekompozicije u zavisnosti od toga da li je model trokomponentni ili četvorokomponentni (prisustvo efekta kalendara, tj. trgovačkih dana,  $TD_t$ ):

	Model sa tri komponente	Model sa četiri komponente
Komponente	$X_t = \text{originalna}$ $T_t = \text{trend}$ $S_t = \text{sezonska}$ $I_t = \text{irregularna}$	$X_t = \text{originalna}$ $T_t = \text{trend}$ $S_t = \text{sezonska}$ $I_t = \text{irregularna}$ $TD_t = \text{trgovacki dani}$
Multiplikativni	$X_t = T_t \cdot S_t \cdot I_t$ $SA_t = X_t/S_t = T_t \cdot I_t$ $S_t, I_t, TD_t$ su centrirani oko jedinice	$X_t = T_t \cdot S_t \cdot TD_t \cdot I_t$ $SA_t = X_t/(S_t \cdot TD_t) = T_t \cdot I_t$
Aditivni	$X_t = T_t + S_t + I_t$ $SA_t = X_t - S_t = T_t + I_t$ $S_t, I_t, TD_t$ su centrirani oko nule	$X_t = T_t + S_t + TD_t + I_t$ $SA_t = X_t - S_t - TD_t = T_t + I_t$
Pseudo-aditivni	$X_t = T_t \cdot (S_t + I_t - 1)$ $SA_t = X_t - T_t \cdot (S_t - 1)$ $S_t, I_t, TD_t$ su centrirani oko jedinice	$X_t = T_t \cdot TD_t \cdot (S_t + I_t - 1)$ $SA_t = X_t/TD_t - T_t \cdot (S_t - 1)$

**Tabela 1:** Tabelarni prikaz najčešćih modela dekompozicije

#### 4.4.7 Direktni/indirektni pristup desezoniranju

Direktno sezonsko prilagođavanje sprovodi se ako se sve vremenske serije (uključujući aggregate)<sup>10</sup> desezoniraju pojedinačno. Na primer, neka su  $\{X_t\}$ ,  $\{Y_t\}$  i  $\{Z_t\}$  tri vremenske serije, pri čemu je  $Z_t = X_t + Y_t$ , za  $\forall t$ . Nakon desezoniranja dobijaju se rezultati za svaku seriju i agregat pojedinačno:  $SA(X_t)$ ,  $SA(Y_t)$ ,  $SA(Z_t)$ , za  $\forall t$ , pri čemu važi:  $SA(Z_t) \neq SA(X_t) + SA(Y_t)$ .

Indirektno sezonsko prilagođavanje sprovodi se ako se procena desezoniranja za vremensku seriju izvodi kombinovanjem procene za dve ili više direktno prilagođenih serija. Na primer, neka su  $\{X_t\}$ ,  $\{Y_t\}$  i  $\{Z_t\}$  tri vremenske serije, gore spomenute. Nakon desezoniranja dobijaju se rezultati za  $\{X_t\}$  i  $\{Y_t\}$  pojedinačno:  $SA(X_t)$  i  $SA(Y_t)$ , za  $\forall t$ , a onda se  $Z_t$  izračunava:  $SA(Z_t) = SA(X_t) + SA(Y_t)$ .

Problem direktnog/indirektnog prilagođavanja je relevantan u različitim slučajevima, npr. unutar procena na nivou sektora ili agregiranja procena sličnih vremenskih serija iz različitih geografskih entiteta. Još uvek je otvoreno pitanje koji je od ova dva pristupa prikladniji. Ni teorijski ni empirijski dokazi ne idu jednoglasno u prilog jednom ili drugom pristupu. Direktni pristup se češće koristi zbog transparentnosti i tačnosti, naročito kada komponentne serije pokazuju slične sezonske obrasce. Indirektni pristup se češće bira kada komponentne serije pokazuju obrasce koji se bitno razlikuju. U svim indirektno desezoniranim agregatima treba uvek proveriti prisustvo rezidualne sezonalnosti. [15, 16]

10. Agregati su sumarni indikatori i ključne veličine za makro-ekonomski analize i poređenja u vremenu i prostoru, koji mere rezultate ukupne ekonomije posmatrane sa različitim stanovišta.

#### 4.4.8 Specifični problemi desezoniranja

- **Desezoniranje kratkih vremenskih serija**

Za neke serije koje su prekratke za desezoniranje (bilo primenom *TRAMO-SEATS*, *X-12* ili *X-13* metode), moguće je prilagođavanje korišćenjem alternativnih, manje standardnih postupaka. Za serije koje su dovoljno duge za obradu pomoću ovih metoda, a opet prilično kratke (3-7 godina), mogu se pojaviti neki problemi oko nestabilnosti. Urađeno je više empirijskih upoređivanja kako bi se istražile relativne osobine ovih metoda na kratkim serijama. Opšte je pravilo kod serija kraćih od sedam godina da specifikacije parametara, korišćenih za prethodni tretman i desezoniranje, treba proveravati češće (npr. dvaput godišnje), kako bi se rešio problem višeg stepena nestabilnosti takvih serija.

U principu, savetuje se da serije kraće od tri godine ne treba sezonski prilagođavati. Sezonsko prilagođavanje kratkih vremenskih serija (3-7 godina) treba sprovoditi pomoću standardnih alata, kadgod je to moguće. Treba sprovoditi simulacije o relativnim karakteristikama postojećih standardnih alata za prilagođavanje kratkih serija. Korisnike treba informisati o većoj nestabilnosti desezoniranih podataka za kratke serije, kao i o primenjenim metodama. Treba definisati jasna pravila politike objavljivanja. Postavke i parametre desezoniranja treba proveravati češće od jednom godišnje.

Alternativa je uopšte ne sprovoditi desezoniranje sasvim kratkih vremenskih serija, a treba izbegavati upotrebu nestandardnih alata za kratke vremenske serije. [15, 24]

- **Tretman problematičnih serija**

Specifične osobine nekih serija su na primer:

- visoka nelinearnost, koja ne dopušta identifikaciju modela s prihvatljivom dijagnostikom modeliranja, čak i skraćivanjem serije;
- odsutnost jasnog signala zbog prisutnosti dominantne iregularne komponente (npr. mali sezonski vrhovi u diferenciranim, logaritmowanim i originalnim podacima);
- nestabilna sezonalnost;
- veliki broj autlajera u odnosu na dužinu serije (tj. više od 10% nepravilnih tačaka);
- heteroskedastičnost (u seriji/komponentama) koja nije ograničena na nekoliko kalendarskih meseci ili koja se ne može izbeći brisanjem podataka iz prethodnih godina, ostavljajući dovoljno podataka za procenu modela.

Ovakve serije ne mogu se desezonirati na standardni način: treba sprovesti *ad hoc* tretman u odnosu na programsku podršku (softver) i skup opcija. Kvalitet sezonski prilagođenih podataka zavisiće od prikladnosti usvojene strategije.

Kod problematičnih serija najčešće se bira pristup „desezoniranje od slučaja do slučaja“ u odnosu na standardni pristup. Kako bi se došlo do rešenja, treba se konsultovati sa literaturom, priručnicima i stručnjacima. Korisnike treba informisati o usvojenoj strategiji. [15, 24]

#### 4.4.9 Prognoza (engl. *Forecasting*)

Jedan od ciljeva izbora modela vremenske serije je prognoziranje njenih budućih vrednosti. Neka je data vremenska serija  $X_1, \dots, X_T$ . Želimo da ocenimo vrednost  $X_{T+h}$ , gde je  $h$  ceo broj, koji se zove intuitivni horizont. Prognozirana vrednost  $X_{T+h}$ , u trenutku  $T$  za  $h$  koraka unapred, obično se označava sa  $\hat{X}_T(h)$ . Prilikom prognoziranja mora se odabrati odgovarajući postupak. Predviđanje je zapravo oblik ekstrapolacije. Prognoze se uvek moraju pripremiti, po potrebi i menjati, ako se pojave dodatne informacije.

Generalna podela metoda za predviđanje je na subjektivne, univarijantne i multivarijantne metode. Prognoze su *subjektivne* ako nastaju po sopstvenom nahođenju. Uglavnom se žele projekcije, koje su najčešće delom objektivne. Prognoze za datu promenljivu pri *univarijantnom* prognoziranju zasnovane su na modelu poznatih serija. Tako je ocena  $\hat{X}_T(h)$  zavisna samo od vrednosti  $X_T, X_{T-1}, X_{T-2}, \dots$ . Na primer, nagoveštaj za buduću prodaju nekog proizvoda temelji se samo na prethodnoj promeni i ne uključuje druge uticaje. Pri *multivarijantnom* predviđanju, prognoze za datu promenljivu bar delimično se zasnivaju na vrednostima jedne ili više dodatnih promenljivih, koje se zovu intuitivne promenljive. Prognoza o prodaji može zavisiti, na primer, od podele ekonomskih pokazatelja. Izbor metode zavisi od podataka koji su na raspolaganju, dužine intuitivnog horizonta, resursa, veština i iskustva analitičara.

Prognoza se može izraziti kao vrednosna tačka ili kao interval u kojem leži buduća vrednost sa određenom verovatnoćom. Mali interval nije nužno dobar, jer interval odražava neizvesnost prognoze.

Prepostavimo da u trenutku  $T$  treba prognozirati vrednost  $X_{T+h}$ , koja je  $h$  koraka ispred. Prediktor za  $X_{T+h}$  se oslanja na skup informacija, koji se osnačava sa  $I_T$ , a sadrži sve dostupne informacije u trenutku  $T$ , koje se mogu koristiti u predviđanju. Obično se pretpostavlja da skup informacija u bilo kom trenutku  $T$  sadrži vrednost  $X_T$  i sve prethodne vrednosti:  $I_T = \{X_{-\infty}, \dots, X_{T-1}, X_T\}$ . Najbolji prediktor za  $X_{T+h}$  označava se sa  $X_{T+h|T}$  i jednak je  $X_{T+h|T} \equiv E(X_{T+h}|I_T)$ . Ovaj prediktor minimizira očekivani kvadrat greške predviđanja.

Kod predviđanja nije važna samo vrednost koja se predviđa, već i tačnost te prognoze. Da bi se procenila tačnost, definiše se greška prognoze  $X_{T+h} - X_{T+h|T}$  i očekivani kvadrat greške predviđanja je:  $c_h \equiv E((X_{T+h} - X_{T+h|T})^2) = \text{Var}(X_{T+h}|I_T)$ . [15, 20]

#### 4.4.10 Revizija

Revizija sezonski prilagođenih podataka vrši se iz dva glavna razloga. Kao prvo, desezonirani podaci mogu biti revidirani zbog revidiranja neprilagođenih (sirovih) podataka. Ova revidiranja neprilagođenih podataka mogu biti rezultat poboljšanog skupa informacija (u odnosu na pokrivenost i/ili pouzdanost). Kao drugo, revidiranje sezonski prilagođenih podataka može se takođe pojaviti zbog bolje procene sezonskog obrasca zahvaljujući novim informacijama koje pružaju novi neprilagođeni podaci, te zbog karakteristika filtera i postupaka za uklanjanje sezonske i kalendarske komponente. Sve dok se revidiranja temelje samo na novim informacijama, ona su većinom korisna. Međutim, kod desezoniranja se može dogoditi da samo jedno dodatno opažanje rezultuje revidiranjem sezonski prilagođenih podataka za više godina, što katkad zbujuje korisnike [15].

Izazov je pronaći ravnotežu između potrebe za najboljim mogućim desezoniranim podacima, posebno na kraju serije, i potrebe za izbegavanjem nevažnih revizija koja se kasnije mogu povući, vratiti u početno stanje (kompromis između preciznosti sezonski prilagođenih podataka i njihove stabilnosti kroz vreme).

Pre nego što se razvije politika revidiranja, potrebno je razmotriti potrebe korisnika i raspoložive resurse za njenu primenu. Važno je da politika revizije bude što je moguće transparentnija, da ne bi dovela do objavljivanja suboptimalnih sezonski prilagođenih podataka, koji bi mogli dovesti u zabludu korisnike pri interpretiranju ekonomskog stanja [15].

Način na koji se sprovodi sezonsko prilagođavanje ima uticaj na reviziju desezoniranih podataka. Postoji nekoliko mogućnosti: [25, 26]

- *tekuće prilagođavanje* (engl. *Current adjustment*) - model, filteri, autlajeri i parametri regresije se reidentifikuju, a odgovarajući parametri i faktori ponovo procenjuju u postavljenim razdobljima za reviziju. Sezonski i kalendarski faktori, koje treba koristiti pri prilagođavanju sezonskih i kalendarskih učinaka za nove originalne podatke između razdoblja revizije, jesu oni koji su procenjeni u prethodnom razdoblju revidiranja i prognozirani do sledećeg razdoblja revidiranja.
- *istovremeno prilagođavanje* (engl. *Concurrent adjustment*) - model, filteri, autlajeri i parametri regresije se reidentifikuju, a odgovarajući parametri se ponovo procenjuju svaki put kada novi ili revidirani podaci postanu raspoloživi.

Strategija tekućeg prilagođavanja svodi na najmanju moguću meru učestalost revizije i koncentriše se na razdoblje revidiranja. Strategija istovremenog prilagođavanja ostvaruje najtačnije sezonski prilagođene podatke u bilo kojoj vremenskoj tački, ali podrazumeva više postupaka revizije, od kojih će mnogi biti mali i možda ne biti dosledni. Obe strategije imaju nedostatke - prva, npr., može dovesti do nedostatka preciznosti u proceni poslednjih prilagođenih podataka, a druga do visoke nestabilnosti sezonskog obrasca. Zbog toga se u praksi koriste druge, uravnotežene alternative između te dve krajnosti, kako bi se savladale specifičnosti podataka, a zadovoljio dobar kvalitet.

- *delimično istovremeno prilagođavanje* (engl. *Partial concurrent adjustment*) - model, filteri, autlajeri i kalendarski regresori se reidentifikuju jednom godišnje, a odgovarajući parametri i faktori se ponovo procenjuju svaki put kad novi ili revidirani podaci postanu dostupni.
- *kontrolisano tekuće prilagođavanje* (engl. *Controlled current adjustment*) - prognozirani sezonski i kalendarski faktori izvedeni iz tekućeg prilagođavanja koriste se za desezoniranje novih ili revidiranih originalnih podataka. Međutim, sprovodi se interna provera u odnosu na rezultate delimičnog istovremenog prilagođavanja, a ona se naročito koristi ako postoji primetna razlika, što znači da svaku seriju treba sezonski prilagoditi dva puta. Taj pristup je praktičan samo za ograničen broj važnih serija.

Potpuni pregled svih parametara desezoniranja treba vršiti bar jednom godišnje i kad god se pojave signifikantna revidiranja.

Po pravilu, kada se ponovo procenjuju sezonski faktori, menjaju se rezultati desezoniranja od početka vremenske serije. Te se promene mogu, ali ne moraju nužno, objaviti u celini. Dva faktora idu u prilog tome da se revizija uvek sprovodi od početka serije: metodološki identičan tretman svih vrednosti i činjenica da je proračun desezoniranih rezultata lako razumeti i ponoviti. Međutim, postavlja se pitanje sadrži li tekući, novododati podatak zaista relevantnu informaciju za signifikantne revidirane procene uobičajenih sezonskih fluktuacija u prethodnim decenijama. Kao način uravnoteženja informacijskog dobitka i odabranog razdoblja, razdoblje revizije za sezonski prilagođene podatke je često u praksi ograničeno.

#### 4.4.11 Kvalitet desezoniranja

Desezoniranje je kompleksan tretman statističkih podataka, koji traži tačan monitoring pre nego što se prihvate rezultati. Kako bi se osigurali kvalitetni sezonski prilagođeni podaci, treba ih vrednovati koristeći se širokim rasponom mera kvaliteta. Između ostalog, treba pažljivo proceniti odsutnost rezidualnih sezonskih i/ili kalendarskih efekata, te stabilnost sezonski prilagođenog obrasca. Vrednovanje sezonski prilagođenih podataka može se sprovesti pomoću više grafičkih, deskriptivnih, neparametarskih i parametarskih kriterijuma, uključenih u izlaz programa za desezoniranje. Po mogućnosti, to se može dopuniti grafičkom dijagnostikom i dodatnim statističkim testovima, dobijenim iz spoljašnjih statističkih paketa.

Preporučuje se korišćenje detaljnog skupa grafičkih, deskriptivnih, neparametarskih i parametarskih kriterijuma za vrednovanje desezoniranja. Treba ponoviti desezoniranje s drugim skupom opcija u slučaju neprihvatanja rezultata. Posebnu pažnju treba posvetiti sledećim poželjnim karakteristikama desezonirane serije: odsutnosti rezidualne sezonalnosti, odsutnosti rezidualnih kalendarskih efekata, odsutnosti prekomernog prilagođavanja (engl. *over-adjustment*) sezonskih i kalendarskih efekata, odsutnosti signifikantne i pozitivne autokorelacije za sezonske pomake (engl. *lags*) u slučajnoj (irregularnoj) komponenti, stabilnosti sezonske komponente. Dodatno, treba proveriti i prikladnost identifikovanog modela korišćenog u celom postupku prilagođavanja, korišćenjem standardne dijagnostike i nekih dodatnih razmatranja. Važno je da broj autolajera treba da bude relativno mali i da nisu koncentrisani oko istog razdoblja u godini.

Ako ceo ovaj postupak zahteva angažovanje mnogo resursa za prilagođavanje velikih količina podataka, kao prihvatljiva alternativa koriste se samo zadati (engl. *default*) kriterijumi, definisani unutar različitih alata. Uvek treba izbegavati nevrednovanje sprovedenog desezoniranja ili korišćenje samo osnovnog skupa grafičke i deskriptivne statistike za vrednovanje desezoniranja.

Programi *TRAMO-SEATS* i *X-12* pružaju širok raspon mera kvaliteta, zajedno sa grafičkom i spektralnom analizom. Mere odražavaju, do izvesnog stepena, različite filozofije na kojima se temelje ovi pristupi. Ipak, obuhvataju i neke zajedničke mere.

Potpuna struktura *TRAMO-SEATS*-a bazirana je na modelu, što povlači da su dijagnostike prikladnosti modela posebno važne, jer se koriste za dobijanje dodatnih uvida u postignuti kvalitet dekompozicije (npr. greška procene za desezoniranu seriju, standardna greška revizije i stopa rasta, signifikantnost sezonalnosti itd.), što se proširuje na prirodan način prema prognozama komponenti. Izlaz *TRAMO-SEATS*-a takođe pruža zbirne tablice, koje sadrže dijagnostiku i mere kvaliteta, iz kojih se izdvajaju ekstremne vrednosti i greške. Tako identifikacija problematičnih serija može biti potpuno automatska.

*X-12-ARIMA* pruža široki skup mera kvaliteta, a struktura metode je takva da je deo za pred-prilagođavanje (engl. *pre-treatment*) u osnovi parametarski, a deo za desezoniranje potpuno neparametarski. Dijagnostičke pokazatelje kvaliteta treba koristiti za procenu rezultata svakog koraka u procesu sezonskog prilagođavanja. U izlazu programa uključena je dijagnostička provera modela regARIMA, sprovedena pomoću različitih analiza rezidualnih odstupanja kod procene modela, te dijagnostička provera sezonskog prilagođavanja. Jedinstvenost izlaza je to što pruža i mnoštvo zbirnih pokazatelja za razmatranje kvaliteta desezoniranja.

Iako *TRAMO-SEATS* i *X-12* pružaju svoje vlastite mere kvaliteta, mnoge od njih su zajedničke, pa je moguće konstruisati široki skup zajedničkih mera kvaliteta. Na tome se

radilo i rad nastavljaju *U.S.Census Bureau*, *Eurostat-u* i *Bank of Spain*. Korak bliže ovoj ideji je razvoj softvera JDemetra+, koji koristi i metod *X-13-ARIMA-SEATS* - objedinjenje gore pomenutih metoda. Upotreba skupa zajedničkih mera kvaliteta korisna je kod upoređivanja sezonski prilagođenih podataka dobijenih pomoću različitih metoda i/ili upoređivanjem različitih strategija.

Važno je da desezonirani podaci budu na odgovarajući način dokumentovani korišćenjem standardnog formata, po mogućnosti u skladu sa smernicama *SDMX (Standard Data and Metadata Exchange)*. Metapodaci o procesu desezoniranja vrlo su korisni, ne samo za razmenu informacija, već i za transparentnost korišćenih opcija, a s tim u vezi i usklađenoj primeni smernica za desezoniranje na evropskom nivou. Stoga, treba ih priložiti pri svakoj novoj objavi podataka i redovno ih ažurirati [15].

SEASONAL ADJUSTMENT METADATA TEMPLATE									
LINK TO GUIDELINES / GLOSSARY									
Group of series : name									
<b>Country/Institution</b>									
<b>Contact Person responsible for SA (not a generic contact point)</b>	<i>name / institution /position phone / e-mail /fax</i>								
<b>GENERAL INFORMATION LINK TO OTHER STRUCTURAL METADATA FOR THE ORIGINAL (RAW) SERIES</b>									
<b>Frequency (*)</b>									
<b>Published series / relevant series</b>	<i>Number of series by level of breakdown and kind of adjustment</i>								
<i>Kind of adjustment</i>	<i>Level of breakdown:</i>								
	<i>Raw series</i>								
	<i>Only calendar adjusted series</i>								
	<i>Other adjustments (e.g. outliers correction)</i>								
	<i>Only seasonally adjusted</i>								
	<i>Seasonally and calendar adjusted,</i>								
	<i>Trend-cycle</i>								
	<i>Others</i>								
<i>(*) if the indicators are published at a quarterly level but are available at a monthly level, please indicate if the adjustment is made at monthly or at quarterly level</i>									
<b>Method used</b>	<i>Parametric (please specify) Non parametric (please specify)</i>								
<b>Software used (please also specify the version)</b>	<i>Software</i>					<i>Version</i>			
<b>Publications</b>	<i>title of hardcopy on-line access source</i>								

Slika 4.12: Deo primera obrasca metapodataka za desezoniranje, [15]

## 5 ALAT ZA DESEZONIRANJE - JDemetra+ (2.0.0)

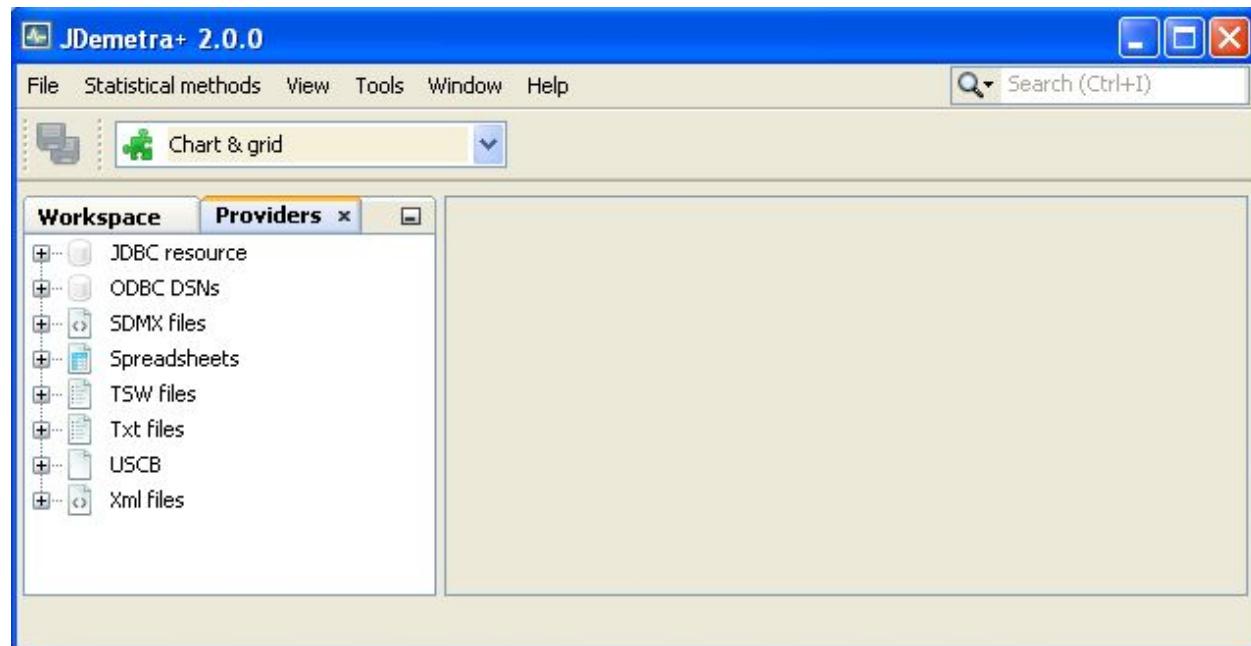
JDemetra+<sup>11</sup> je *open source* softver sa nezavisnom platformom i mogućnošću za nadogradjivanje, čija je primarna namena desezoniranje vremenskih serija, a koji je razvila Nacionalna banka Belgije. Sadrži najnovije verzije vodećih algoritama za desezoniranje, koji prate metodologiju objektno orijentisanog programiranja. JDemetra+ ne samo da ima *user friendly* grafički interfejs, uporediv sa svojim prethodnikom (Demetra+), već sadrži i skup raspoloživih Java biblioteka, koje mogu da se koriste i za neke druge, slične potrebe. Izgrađena je na konceptu i algoritmima, koji koriste dve vodeće SA<sup>12</sup> metode - TRAMO/SEATS i X-12-ARIMA/X-13ARIMA-SEATS.

Najnovija verzija, JDemetra+ 2.0.0, dostupna je od januara 2015. godine i sve analize u radu izvršene su njenom upotrebljom. Rezultati koje ovaj alat nudi za istu seriju razlikuju se u zavisnosti od izabranih statističkih/algoritamskih izbora, no opšta poruka o desezoniranju je gotovo uvek slična. Jedan od ciljeva razvojnog procesa JDemetra+ je omogućavanje upoređivanja rezultata između metoda TRAMO/SEATS i X-12-ARIMA/X-13ARIMA-SEATS. Iz ovog razloga, većina alata za analizu je slična u oba algoritma, što nije bio slučaj u prethodnim verzijama.

Većina primera u ovom radu biće konceptualno vezana za metod X-12-ARIMA/X-13ARIMA-SEATS, ali će biti predstavljena i razlika u rezultatima koje nude obe metode, a koja može nastati pri određenim karakteristikama vremenske serije.

### 5.1 Pregled JDemetra+

Nakon pokretanja programa, klikom na ikonicu  , aktivira se osnovni (*default*) prozor:



Slika 5.1: Osnovni prozor programa

11. Najnovija verzija dostupna je korisnicima na: <https://github.com/jdemetra/jdemetra-app/releases>

12. SA je skraćenica od engleske reči *Seasonal Adjustment* (prevod: desezoniranje)

Najpre je preporučljivo sačuvati *Workspace* pod prepoznatljivim imenom i na lokaciju koju je lako ponovo naći. Opcija za čuvanje je, očekivano, *File → Save Workspace As...*. Sačuvani program može se ponovo otvoriti, korigovati i modifikovati u skladu sa novopristiglim ili izmenjenim indeksima u seriji. Treba napomenuti da je to jedino moguće ako lokacija i ime ne podležu izmenama.

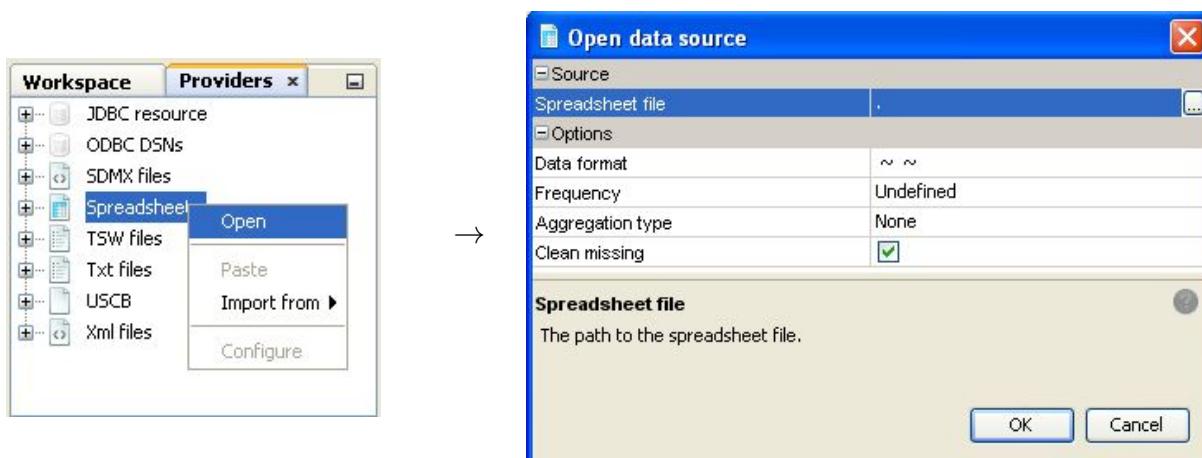
## 5.2 Importovanje serije/grupe serija

Aktivirani *Providers* panel (sa slike 5.1) služi za importovanje serija (sa definisanim lokacija na lokalnom računaru), koje će biti predmet nekog vira obrade. Mogući izvori podataka izlistani su na panelu, a najčešće korišćeni među njima je *Spreadsheets*, s obzirom na to da odgovara serijama pripremljenim u Excel fajlu. JDemetra+ zahteva da ulazni fajl zadovoljava određene tehničke kriterijume: prvi red rezervisan je za nazine i/ili šifre serija, a prva kolona za datume zabeleženih observacija. Gornja leva ćelija (A1) može, ali ne mora biti popunjena bilo kakvim tekstrom. Prazne ćelije JDemetra+ interpretira kao vrednosti koje nedostaju (engl. *missing value*) i one se mogu pojaviti na početku, u sredini i na kraju vremenske serije.

A	B	C	D	E	F	G
1 date	7030	7040	7050	7060	9022	9023
2 1-1999	2.7	36.2	3	41.4	1.3	3.4
3 2-1999	2.8	37	3.2	42.3	1.3	3.4
4 3-1999	2.9	38.3	3.3	43.9	1.5	3.7
5 4-1999	2.5	31.3	2.8	35.7	1.2	2.8
6 5-1999	2.2	27	2.5	30.9	1.1	2.7
7 6-1999	2.2	26.9	2.5	30.8	1.2	3
8 7-1999	2.6	31.1	2.9	35.6	1.5	3.6
9 8-1999	2.7	31.6	3.1	36	1.5	3.1
10 9-1999	3	31.9	3.5	36.5	1.3	3.1
11 10-1999	3.3	31.9	3.8	36.5	1.7	4.7
12 11-1999	3.6	33.8	4.1	38.7	1.7	4.5
13 12-1999	4	37.3	4.6	42.7	2.3	6.6

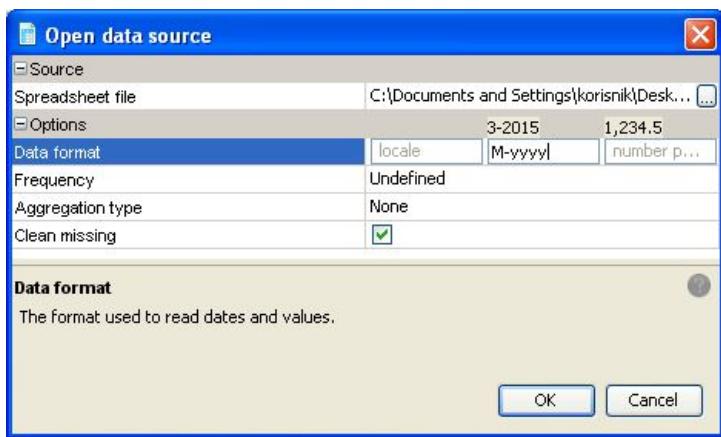
Slika 5.2: Primer pravilno definisanog ulaznog Excel fajla

Postupak učitavanja željenog Excel fajla započinje se aktiviranjem *Spreadsheets* opcije u *Providers* panelu, pri čemu se otvara novi panel, *Open data source*:



Slika 5.3: Postupak učitavanja Excel fajla

Ikonica u polju *Spreadsheet file* nudi mogućnost lociranja želenog fajla, polje *Data format* definisanje formata prve kolone fajla koji se učitava, *Frequency* definisanje frekven-cije observacija (mesečne, kvartalne, polugodišnje, godišnje..., pri čemu opcija *Undefined* prepoznaće mesečne podatke), *Aggregation type* nudi mogućnost odabira tipa aggre-gacije (suma, prosek, minimum, maksimum) ukoliko je to za neku obradu neophodno, a čekirana opcija *Clean missing* automatski 'čisti' vrednosti koje u Excel fajlu nedostaju.



Slika 5.4: Opcije za unos fajla sa slike 5.2

Nakon učitavanja podataka, ako se u stablu izlista *Spreadsheets*, uočava se staza do foldera u kome je smešten excel fajl, naziv *sheet-a* tog excel fajla i nazivi serija u tom *sheet-u*. Desnim klikom na stazu otvaraju se mogućnosti za editovanje, kloniranje, zatvaranje, preimenovanje, kopiranje i slično. Dvostrukim klikom na bilo koju izlistanu seriju, dobija se njen grafik (engl. *Chart*). Desni klik izlistava opcije za manipulisanje datim grafikom, a prevlačenjem miša po njegovoj liniji prikazuju se vrednosti i datumi zabeleženih podataka. Osim grafičkog prikaza, moguće je videti i matricu sa podacima (engl. *Grid*).

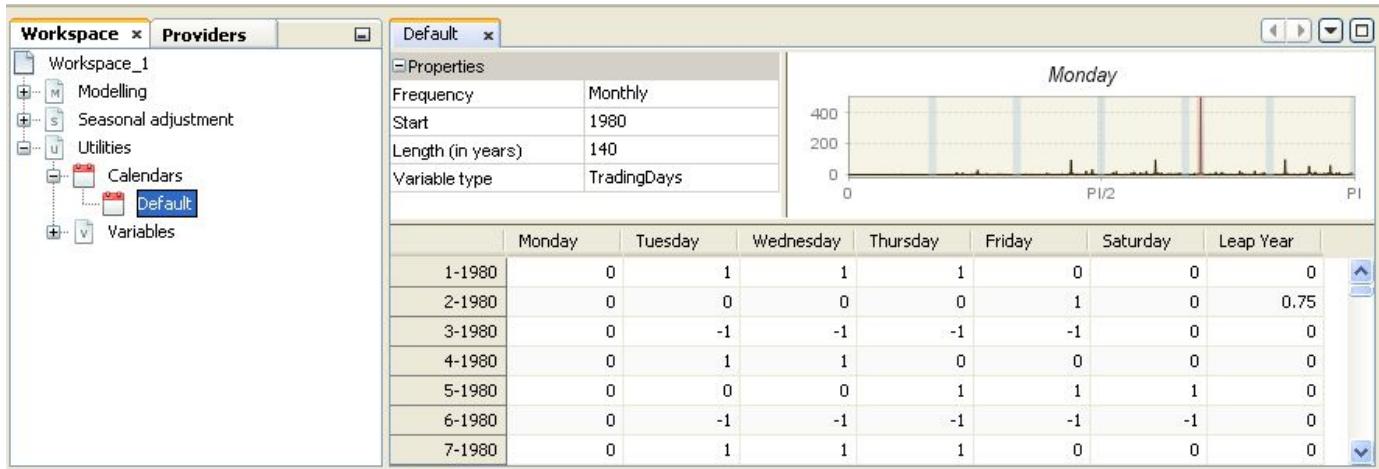


Slika 5.5: Prikaz odabrane serije

### 5.3 Kreiranje i importovanje nacionalnog kalendarra

Za otkrivanje i ocenjivanje efekta kalendara neophodno je definisati nacionalni kalendar, koji će sadržati praznike specifične za određenu zemlju. JDemetra+ sadrži *Default* kalendar, koji kao neradne dane računa samo subote i nedelje, te reflektuje uobičajenu kompoziciju sedmica u periodu koji se posmatra. Analiza je svakako preciznija ako se serije tretiraju u skladu sa svim efektima koji su na njihove vrednosti uticale. Zato se uvek preporučuje definisanje nacionalnih praznika.

U panelu *Workspace*, u odeljku *Utilities* nalazi se izlistan *Default* kalendar. Dvostrukim klikom otvaraju se tri panela, koja sadrže opcije za podešavanje (engl. *Properties*), zatim promenljive kalendara i spektralni graf za odabranu (po *default*-u pondeljak) promenljivu.



Slika 5.6: Osnovni prikaz opštег kalendara (*trading days*)

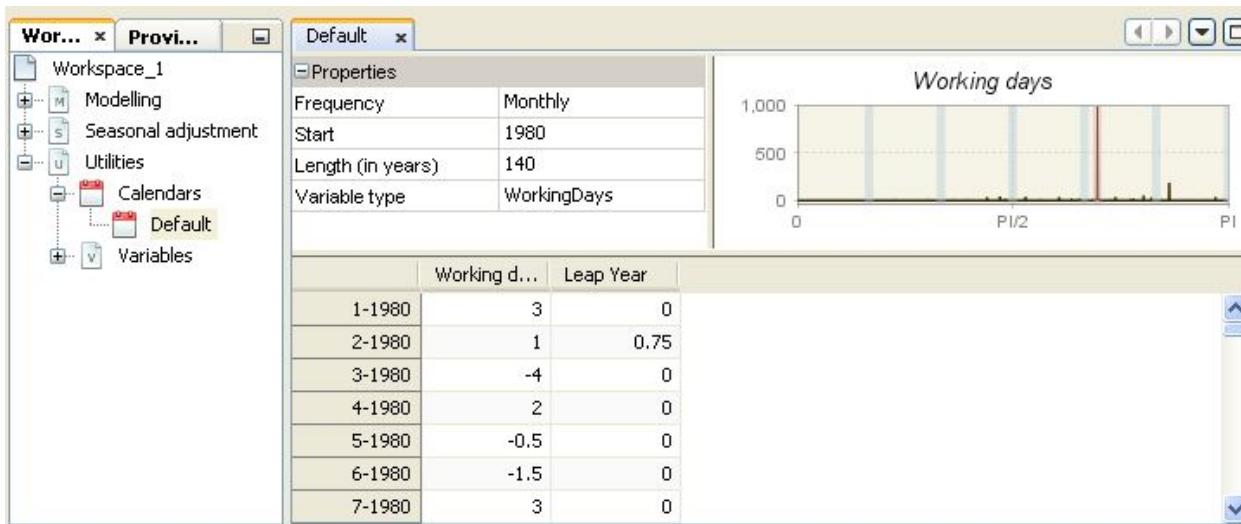
U *Properties* panelu korisnik može da definiše različitu učestalost podataka, početnu godinu, dužinu serije u godinama i tip promenljive - trgovacke/radne dane (engl. *trading/working days*). Za *trading days* opciju kreira se sedam regresora, koji odgovaraju razlikama u ekonomskim aktivnostima između svih dana u nedelji, a uzima u obzir i prestupnu godinu. Uobičajeno je da se ove promenljive dobijaju na sledeći način: *broj pondeljaka - broj nedelja (u mesecu)*; *broj utoraka - broj nedelja (u mesecu)*; *broj sreda - broj nedelja (u mesecu)*..., odnosno, korišćenjem sledećih transformacija:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \#\text{ponedeljaka} \\ \#\text{utoraka} \\ \#\text{sreda} \\ \#\text{četvrtaka} \\ \#\text{petaka} \\ \#\text{subota} \\ \#\text{nedelja} \end{bmatrix} = \begin{bmatrix} \#\text{ponedeljaka} - \#\text{nedelja} \\ \#\text{utoraka} - \#\text{nedelja} \\ \#\text{sreda} - \#\text{nedelja} \\ \#\text{četvrtaka} - \#\text{nedelja} \\ \#\text{petaka} - \#\text{nedelja} \\ \#\text{subota} - \#\text{nedelja} \\ \text{dužina perioda (u danima)} \end{bmatrix}$$

Ovih šest regresora uzimaju neku od vrednosti iz skupa  $\{-1, 0, 1\}$ . Poslednji regresor oslikava prisustvo prestupne godine i on ima efekta samo na februar, a uzima vrednost 0,75 ako godina jeste prestupna ili -0,25 ako nije. Za ostale mesece uvek je 0.

Za *working days* opciju kreiraju se dva regresora, koji odgovaraju razlici u ekonomskim aktivnostima između radnih dana (od pondeljka do petka) i neradnih dana (subota i nedelja), a računa se i efekat prestupne godine. Drugi regresor se računa isto kao u slučaju opcije *trading days*, a transformacija za računanje prvog regresora je sledeća:

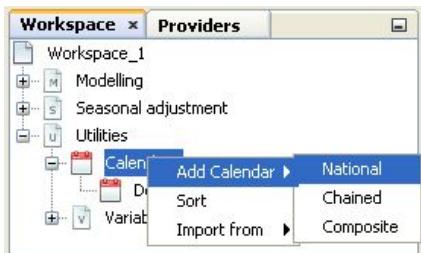
$$\begin{bmatrix} 1 & -5/2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \#\text{radnih dana} \\ \#\text{neradnih dana} \end{bmatrix} = \begin{bmatrix} \text{WD regresor} \\ \text{dužina perioda (u danima)} \end{bmatrix}$$



Slika 5.7: Osnovni prikaz opšteg kalendarja (working days)

Vrhovi na spektru (grafik u gornjem desnom uglu) ukazuju na učestalost cikličnih pomeraja sa periodičnošću manjom od dva meseca. Za mesečne serije sezonske periodičnosti su:  $\frac{\pi}{6}, \frac{\pi}{3}, \frac{\pi}{2}, \frac{2\pi}{3}, \frac{5\pi}{6}, \pi$ . Šiljak na nultoj frekvenciji uvek odgovara trend komponenti serije. Sezonske frekvencije su označene kao siva vertikalna linija, a ljubičaste predstavljaju *trading day* frekvencije. Za kalendarske promenljive se ne očekuje da imaju šiljak (engl. *peak*) na nultoj, niti na sezonskim frekvencijama.

Sve ove promenljive mogu se eksportovati u excel fajl.



Slika 5.8: Opcije za kreiranje kalendarja

Kompozitni kalendar je kombinacija nekoliko nacionalnih kalendarata, sa akcentom na težinskim koeficijentima. Koristan je za serije koje uključuju podatke iz više država ili regiona (na primer, kalendar Evropske unije ili kalendar Nemačke, koja je podeljena na federacije). Okovani (engl. *chained*) kalendar koristi se kada se dogode velike promene u kompoziciji definisanih praznika. Na primer, Poljska je uključila nove praznike 1990. godine, što će se prevazići definisanjem okovanog kalendarata, tj. dva nacionalna kalendarata sa naznačenim datumom prekida (engl. *break date*).

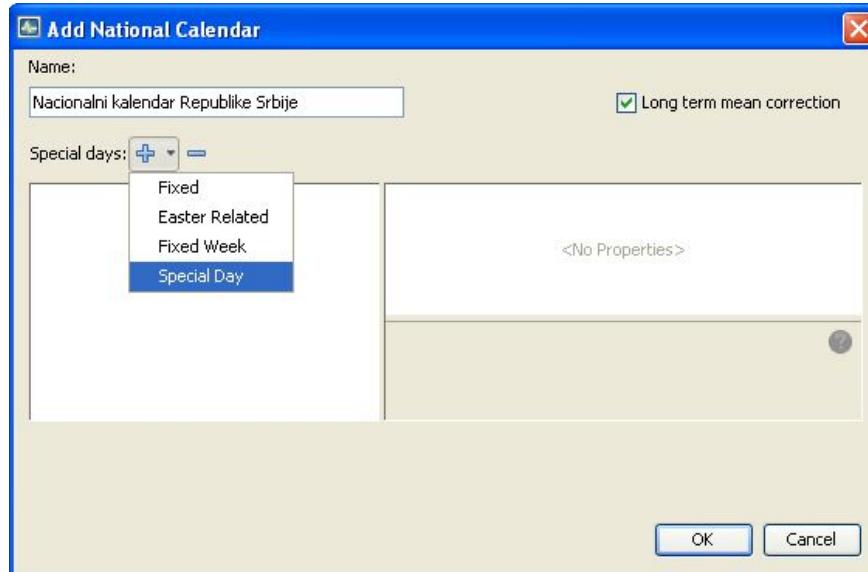
Republika Srbija ima jedinstven kalendar za celu teritoriju, te je za desezoniranje naših serija dovoljno definisati nacionalni kalendar, odnosno uključiti sve praznike, koji se zvanično računaju kao neradni dani. Stoga je najviše pažnje u ovom odeljku posvećeno kreiranju upravo nacionalnog kalendarata.

Na slici 5.9 prikazan je osnovni prozor za kreiranje nacionalnog kalendarata. Najpre treba uneti ime u polju *Name*. Za dodavanje praznika treba izlistati meni , koji daje sledeće opcije:

- fiksni praznici (engl. *Fixed*) - definiše praznike kao specifične dane u godini, a koji su uvek istog dana u mesecu;

Da bi se serije prilagodile efektima specifičnih praznika, preporučuje se kreiranje odgovarajućeg kalendarata. JDemetra+ nudi mogućnost definisanja nacionalnog, kompozitnog i/ili tzv. okovanog (engl. *chained*) kalendarata.

- praznici u relaciji sa Uskrsom (engl. *Easter Related*) - ukazuje na praznike čiji datumi zavise od Uskrsa;
- fiksna nedelja (engl. *Fixed week*) - opcija kojom se kreira fiksni praznik, koji uvek pada u specifičnoj nedelji specifičnog meseca;
- specijalni dani (engl. *Special Day*) - omogućava da korisnik odabere praznik sa već definisane liste, koja uključuje najpopularnije pokretne i fiksne praznike.



Slika 5.9: Kreiranje nacionalnog kalendara - osnovni prozor

Kad se neka od ponuđenih opcija odabere, u desnom prozoru ovog panela otvaraju se opcije za konkretno definisanje praznika:

- *Start* - datum od kog korisnik želi da se definisani praznik računa; ako polje ostane prazno, računa se datum od kog je kalendar definisan;
- *End* - datum do kog korisnik želi da se definisani praznik računa; ako polje ostane prazno, računa se datum do kog je kalendar definisan;
- *Weight* - težinski koeficijent, koji određuje koliki je učinak ovog dana kao neradnog (koliki procenat ljudi ga smatra praznikom); uzima vrednosti između 0 i 1;
- *Day Event* - lista predefinisanih praznika;
- *Offset* - korisnik ima mogućnost da modifikuje poziciju odabranog predefinisanog praznika; to je razlika, u danima, između datuma želenog i predefinisanog praznika; može biti i pozitivan i negativan broj;
- *Day* - dan u mesecu kada se fiksirani praznik slavi;
- *Month* - mesec u kom se fiksirani praznik slavi.

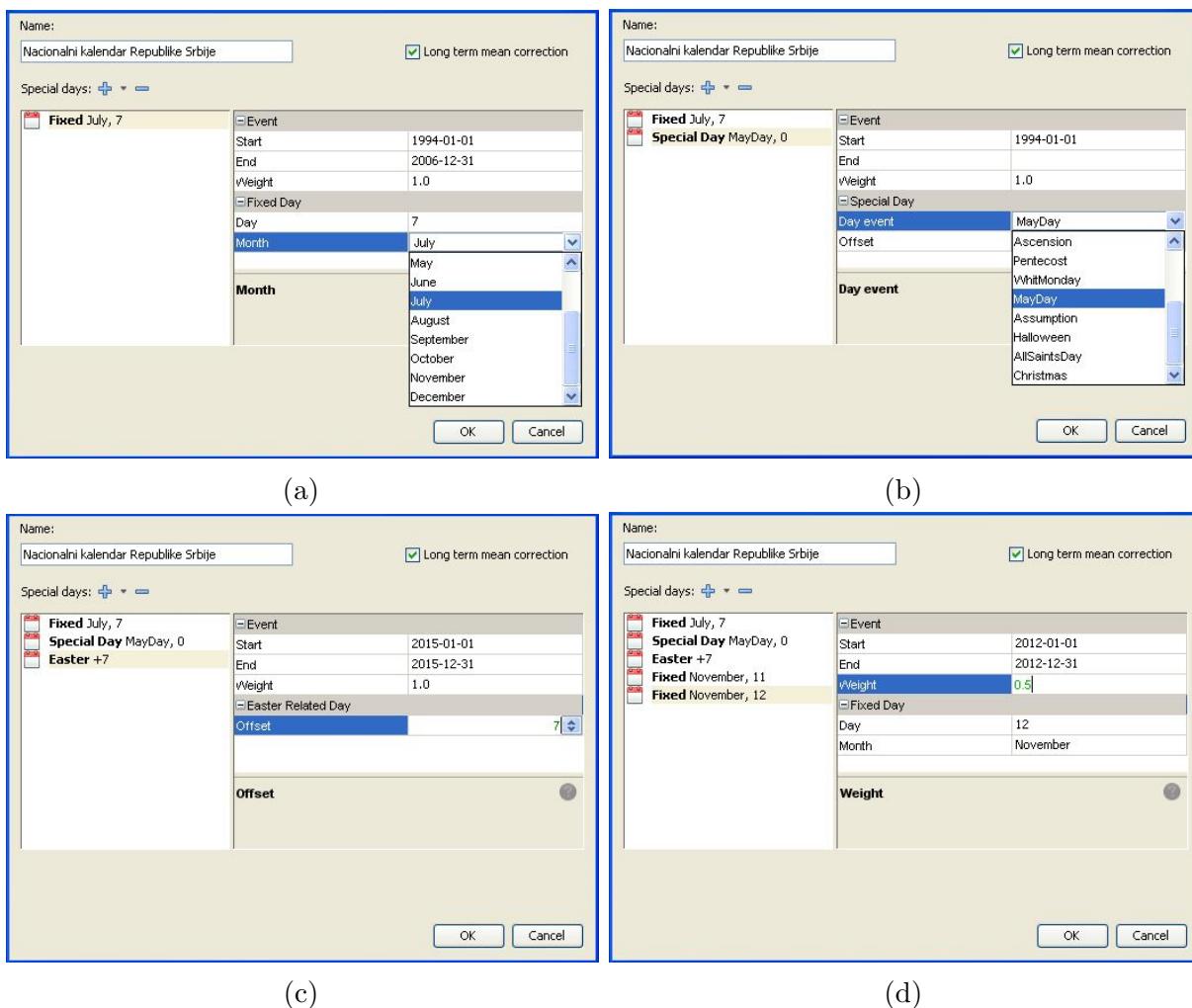
Na slici 5.10 demonstrirani su unosi specifičnih praznika.

Na prvom primeru, pod (a), prikazan je unos Dana ustanka naroda Srbije (7. jul), koji se slavio kao državni praznik do 2006. godine. U skladu sa tim je i polje *End* popunjeno. Početna godina je 1994., jer serije Republike Srbije, koje će se obrađivati, najdalje datiraju od te godine.

Slika (b) predstavlja unos Dana rada (1. maj), koji je jedan od izlistanih specijalnih dana. Polje *End* u ovom slučaju nije popunjeno, jer ne treba da bude ograničeno, a *Start* je iz istog razloga kao i u primeru pod (a) 01.01.1994. godine.

Na trećoj slici, (c), demonstriran je unos pravoslavnog Uskrsa za 2015. godinu. S obzirom na to da JDemetra+ podržava Gregorijanski kalendar, mora se svake godine voditi računa o razlici u danima između katoličkog i pravoslavnog Uskrsa. Ove godine je 7 dana kasnije, pa je u skladu sa tim i definisan *Easter Related Day* sa opcijom *Offset= 7*.

Poslednji primer, pod (d), ima za zadatak da ukaže na mogućnost promene težinskog koeficijenta u slučaju kad je to neophodno. Na primer, Dan primirja u Prvom svetskom ratu, 11. novembar, slavi se od 2012. godine. Te prve godine, pao je u nedelju, koja se već računa kao neradni dan. Stoga, određene državne institucije i privredni sektori praznuju i u ponedeljak, 12. novembra, te je tog dana težinski koeficijent nerada 0.5.

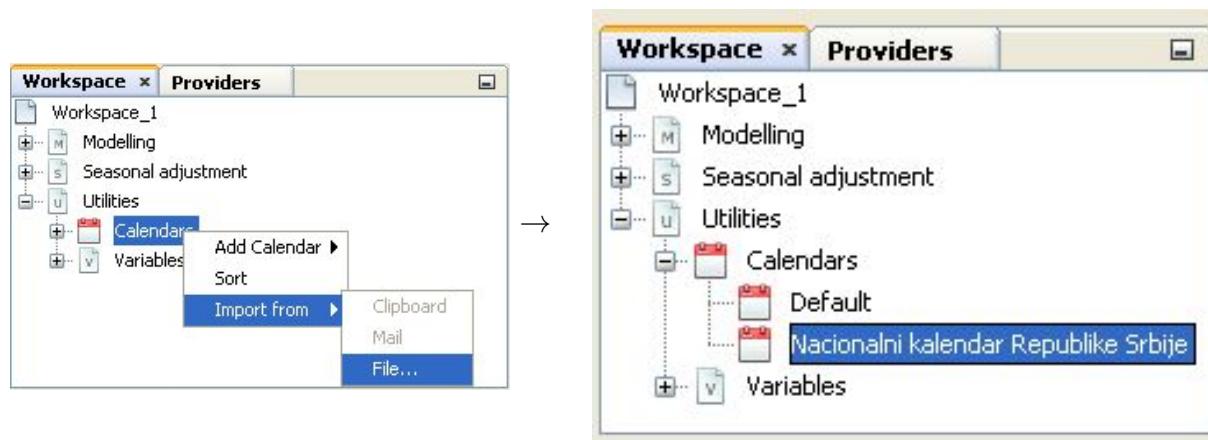


Slika 5.10: Demonstracija unosa pojedinačnih praznika

S obzirom na to da serije Republike Srbije datiraju najdalje od 1994. godine, svaki se praznik od te godine mora uredno uneti, vodeći računa o težinskim koeficijentima, razlici između Julijanskog i Gregorijanskog kalendarja, 'prenošenim' praznicima (ako padaju vikendom), kao i ukinutim i novim državnim praznicima. Takođe, treba uključiti i uticaj najvećih srpskih slava, poput Nikoljdana i Aranđelovdana, i u zavisnosti od učestalosti definisati težinske koeficijente. Početkom svake godine se kalendar mора ažurirati.

U zavisnosti od sektora za čije se serije kalendar koristi, trebalo bi ga modifikovati, ukoliko to priroda poslovanja sektora zahteva. Na primer, u trgovini je mnogo manje neradnih dana, nego u ostalim privrednim granama.

Zbog kompleksnosti, kalendar se ne mora svaki put naknadno definisati. Umesto toga, postoji opcija za importovanje već kreiranog kalendar, nakon čega je on izlistan u stablu *Worspace-a*, kao što je prikazano na slici 5.11. U jednoj obradi može se istovremeno učitati više kalendara za potrebe desezoniranja serija iz različitih privrednih sektora.



Slika 5.11: Postupak učitavanja ranije kreiranog kalendara

## 5.4 Kreiranje specifikacije za proces desezoniranja

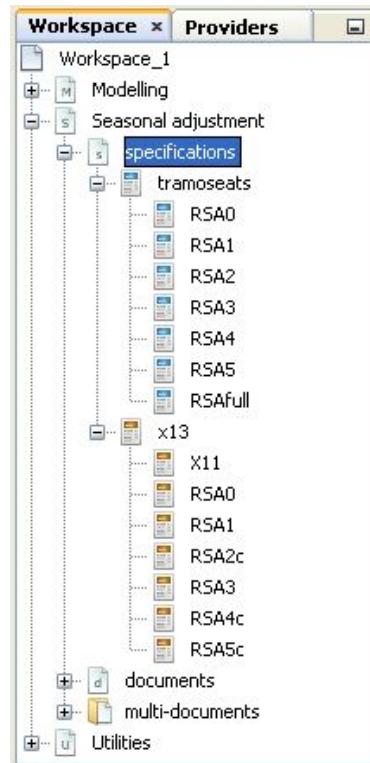
Pre započinjanja procesa desezoniranja, neophodno je odabratи već postojeću specifikaciju (ponuđenu alatom), ili kreirati sopstvenu, u skladu sa potrebama i karakteristikama serija za obradu. U panelu *Workspace*, pod *Seasonal adjustment*, nalazi se nova grana - *Specification*, koja nudi *tramoseats* ili *x13* specifikacije. Odabir vršimo prvenstveno u skladu sa metodom koji želimo da koristimo za obradu. Ove specifikacije su fiksirane i na njima se ne mogu vršiti izmene bilo kakvog tipa.

Opis ponuđenih specifikacija:

- za *tramoseats*:
  - *RSA0* - vazduhoplovni model  $(0, 1, 1)(0, 1, 1)$ ;
  - *RSA1* - log test, autlajeri, vazduhoplovni model;
  - *RSA2* - log test, test za prisustvo efekta *working* dana, efekat Uskrsa, autlajeri, vazduhoplovni model;
  - *RSA3* - log test, autlajeri, automatska identifikacija modela;
  - *RSA4* - log test, test za prisustvo efekta *working* dana, efekat Uskrsa, autlajeri, automatska identifikacija modela;
  - *RSA5* - log test, test za prisustvo efekta *trading* dana, efekat Uskrsa, autlajeri, automatska identifikacija modela;
  - *RSAfull* - log test, efekat *trading/working* dana testiran F-testom, efekat Uskrsa, autlajeri, automatska identifikacija modela;

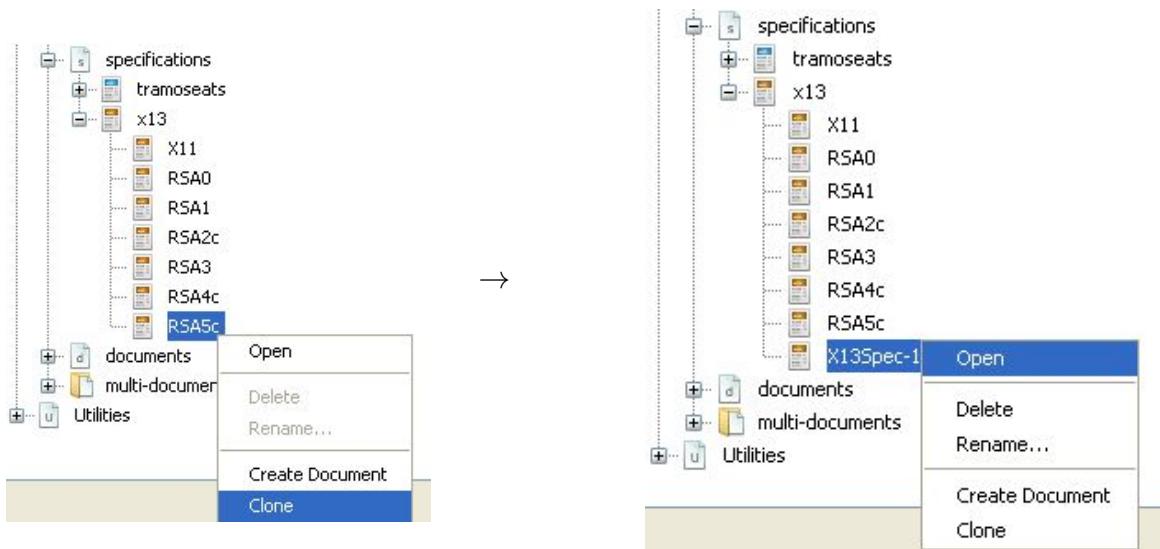
- za x13:

- X11 - bez ikakvog pre-podešavanja;
- RSA0 - vazduhoplovni model  $(0, 1, 1)(0, 1, 1)$ ;
- RSA1 - log test, autlajeri, vazduhoplovni model;
- RSA2c - log test, test za prisustvo efekta *working* dana, efekat Uskrsa, autlajeri, vazduhoplovni model, pre-podešavanje za prestupnu godinu ako je korišćena logaritamska transformacija;
- RSA3 - log test, autlajeri, automatska identifikacija modela;
- RSA4c - log test, test za prisustvo efekta *working* dana, efekat Uskrsa, autlajeri, automatska identifikacija modela, pre-podešavanje za prestupnu godinu ako je korišćena logaritamska transformacija;
- RSA5c - log test, test za prisustvo efekta *trading* dana, efekat Uskrsa, autlajeri, automatska identifikacija modela, pre-podešavanje za prestupnu godinu ako je korišćena logaritamska transformacija;



Slika 5.12: Ugrađene specifikacije

Većina serija Republike Srbije je, pre korišćenja alata JDemetra+, desezonirana X-11 programom, pa kako je X-13 njegov naslednik, uglavnom ima prioritet u odnosu na *tramoseats*. No, s obzirom na to da je analiza serija daleko preciznija ako se uključi efekat nacionalnog kalendara, uvek se preporučuje kreiranje sopstvene specifikacije na bazi neke od ponuđenih, jer se tek tada mogu vršiti izmene. Postupak zahteva kloniranje već definisane specifikacije (koja najbolje odgovara potrebama), uz naknadno menjanje opcija, u skladu sa karakteristikama serija i željene obrade.



Slika 5.13: Kloniranje postojeće specifikacije



Slika 5.14: Opcije kreirane specifikacije

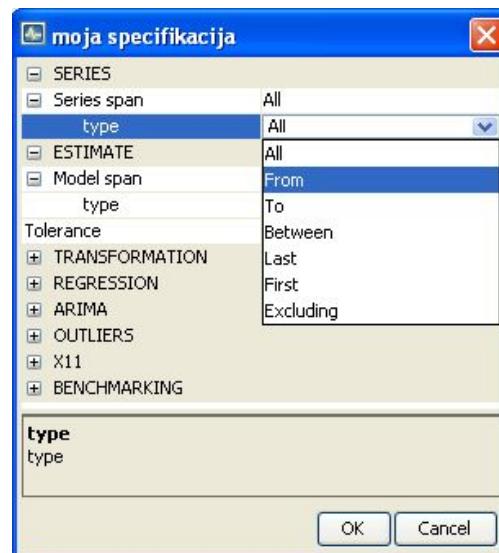
Ovaj prozor nosi naziv "moja specifikacija", jer je *X13Spec-1* ručno preimenovano. Kao što se vidi na slici 5.14, prozor je podeljen u nekoliko sekcija. Konkretan sadržaj zavisi od specifikacije koja je klonirana, ali se obično uzima *RSA5c* za *x13* ili *RSAfull* za *tramoseats*, jer su najkompleksnije. Ove specifikacije, koje korisnik definiše, korisne su za desezoniranje velikih grupa vremenskih serija. Međutim, ukoliko ista specifikacija ne treba da važi za sve serije uključene u proces, izmene se mogu vršiti i na pojedinačnim serijama, nakon aktiviranja procesa desezoniranja.

#### 5.4.1 Skraćivanje serije (engl. *Model/Series span*)

Po *default-u*, JDemetra+ vrši analizu na kompletnoj vremenskoj seriji. Međutim, u nekim slučajevima potrebno je ograničiti analizu na podskup, tj. deo (engl. *span*) vremenske serije. Na primer, veoma duga serija imaće podatke koji se neće slagati sa kretanjem tekućih vrednosti. U kontekstu desezoniranja, moguće je heuristički prepostaviti da su duge serije dužine preko dvadeset godina, a na njima je svakako teško primeniti metod desezoniranja. Stoga je nekad razumno ograničiti vremensku seriju na skorije observacije. Takođe, u slučaju kad je prisutna pojava vrhova (engl. *peaks*) u spektru desezonirane serije ili irregularne komponente, preporučuje se skraćivanje. JDemetra+ nudi dve korisne opcije za rešavanje problema dužine serija:

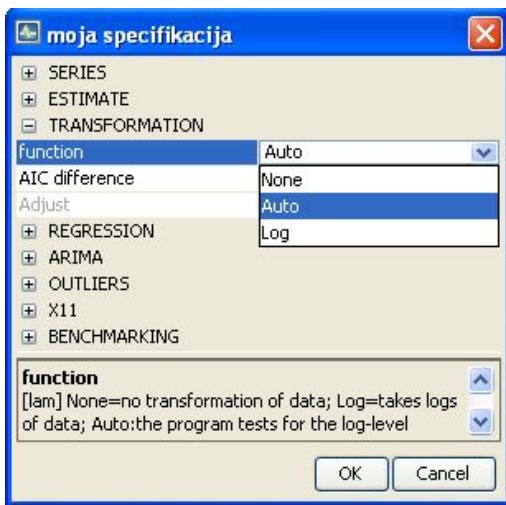
- *Model span* - određuje podskup vremenske serije koji će se koristiti za proces modeliranja/desezoniranja; samo ograničeni deo ulazi u proračune;
- *Series span* - određuje deo koji će se koristiti za ocenu modela u predfazi desezoniranja; koristi se kad, na primer, korisnik ne želi da raniji podaci oštete prognozu.

Deo koji je određen u *model span* opciji koristi se za modeliranje i dekompoziciju rezultata iz podešenog dela u *series span* opciji.



Slika 5.15: Opcije *Model span* i *Series span*

#### 5.4.2 Transformacija serije

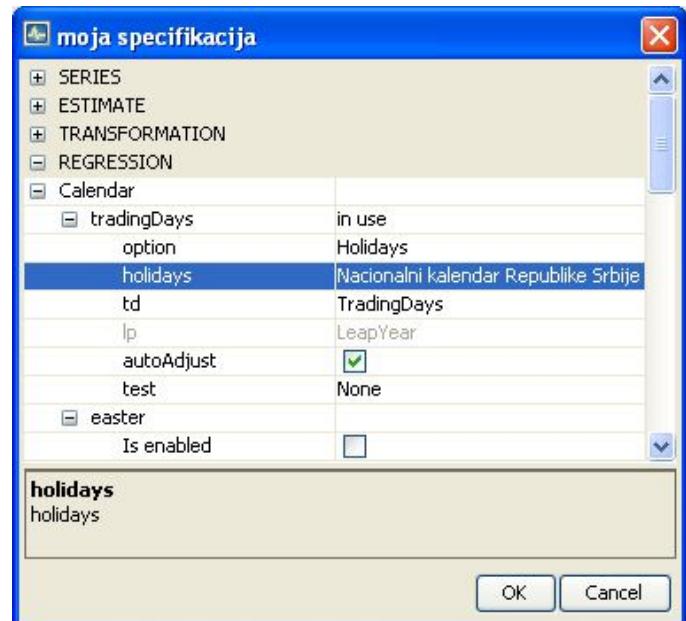


Slika 5.16: Transformacija serije

Logaritamska transformacija originalnih podataka se primjenjuje za postizanje stacionarnosti autokovarijacione funkcije. Odluka o eventualnom primjenjivanju transformacije ima ogroman uticaj na rezultat desezoniranja. JDemetra+ nudi dve opcije: *log* transformacija, što znači da se koristi multiplikativna dekompozicija; i *none* transformacija, što znači da se koristi aditivna dekompozicija. Takođe, postoji opcija *Auto*, kojom korisnik dopušta da softver odluči koja od dve opcije je poželjnija, koristeći se *log-level* testom. Test koji se koristi kod *TramoSeats*-a zasnovan je na oceni parametra po metodi maksimalne verodostojnosti u *Box-Cox* transformaciji, dok je test koji se koristi kod *X-13Arma-Seats*-a zasnovan na Akaikeovom (AICC) informacionom kriterijumu.

#### 5.4.3 Kalendar

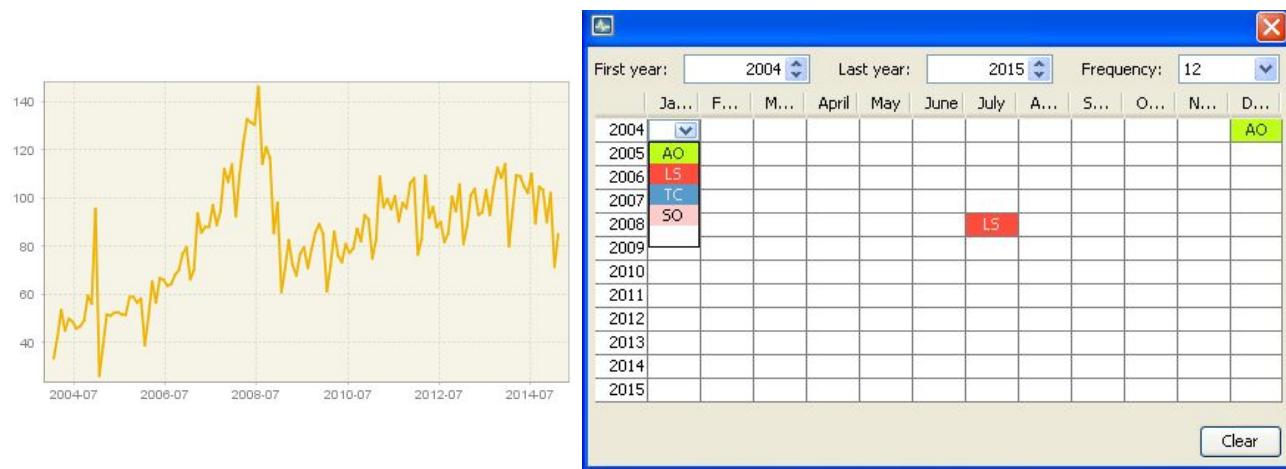
Korisniku je u ovom koraku dozvoljeno da uključi kreirani ili importovani kalendar, umesto korišćenja *default* kalendar, koji nudi program. Takođe, ovde se vrši odabir i između *trading/working* dana, kao i efekta prestupne godine. Opcijom *test* korisnik definiše želi li da program svaki put testira ima li efekat kalendara uticaj na pojedinačnu seriju i treba li da se računa, ili želi da ovaj efekat bude uvek uzet u obzir. Za srpski kalendar opcija *easter* treba da je dečekirana, s obzirom na to da se u nacionalnom kalendaru ručno definiše pravoslavni Uskrs, pa je automatsko računanje i katoličkog višak. Nakon grupne obrade, korišćenjem *Joint F* testa može se pojedinačno proveriti da li strukturi podataka više odgovaraju *trading* ili *working* dani.



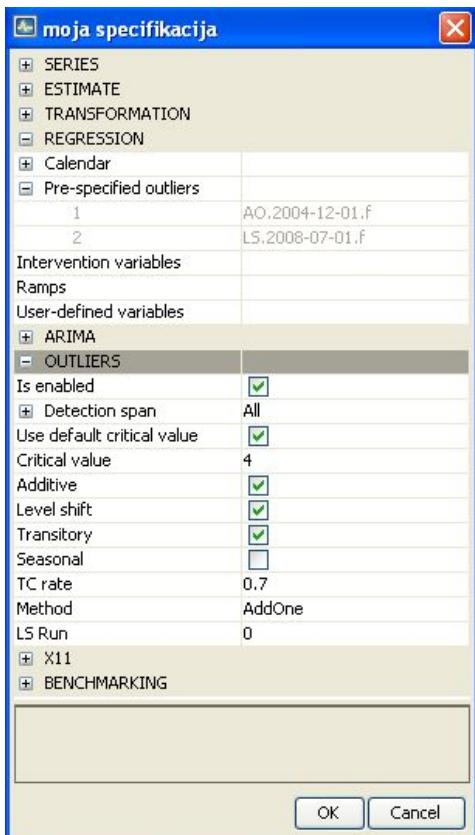
Slika 5.17: Definisanje kalendara

#### 5.4.4 Autlajeri

Vrednosti koje ekstremno odstupaju od 'pravila' kretanja serije, tzv. autlajeri, ugrožavaju modeliranje vremenske serije, jer su u vezi sa iregularnim uticajima na seriju. Takođe, imaju i štetne efekte na kvalitet desezoniranja, jer dovode do pristrasnosti u oceni parametara, siromašne prognoze i neadekvatne dekompozicije serije. Stoga je od izuzetne važnosti da se identifikuju, uključe u korak modeliranja, a njihovi efekti uklone pre dekomponovanja. I *X-13ARIMA-SEATS* i *TRAMO/SEATS* imaju automatske procedure za detekciju i korekciju autlajera. Međutim, a priori informacije o događajima koji mogu da uzrokuju abnormalne observacije korisnik može ručno da unese u model.



Slika 5.18: Ručno unošenje autlajera za datu seriju



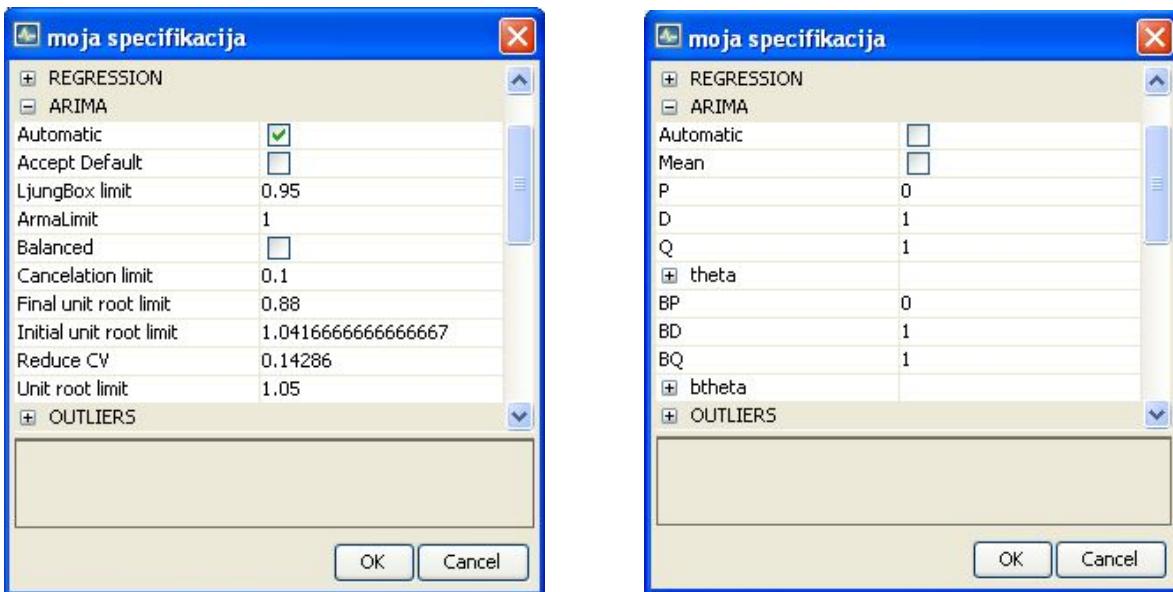
Slika 5.19: Autlajeri

U automatskoj proceduri otkrivanja i uklanjanja autlajera, tri tipa su osnovna: aditivni (AO), *transitory change* (TC) i *level shift* (LS). Neposredno sa grafika originalne funkcije, prikazanog na slici 5.18, mogu se uočiti dve ekstremne vrednosti: u decembru 2004. aditivni i u julu 2008. LS autlajer. U polju *Pre-specified outliers* korisniku je dozvoljeno da ovakva zapažanja ručno zabeleži, naročito ako postoje ekonomski ili socijalna objašnjenja, unapred poznata, za takve iznenadne promene. Po zatvaranju panela u kom nam je dozvoljeno da učitamo autlajere, oni postaju vidljivi u prozoru 'moja specifikacija' (slika 5.19). Slovo 'f' ukazuje da su ovi autlajeri fiksni, što znači da njihova vremenska lokalizacija i tip neće biti verifikovani u procesu desezoniranja. Nakon pokretanja tog procesa, u delu gde je analizirana cela obrada, može se utvrditi da li su ovi unapred definisani autlajeri statistički značajni ili ne, pa se u odnosu na to mogu vršiti izmene i popravke.

S druge strane, automatska identifikacija autlajera takođe je izbor korisnika. Čekiranjem/dečekiranjem ponuđenih tipova ekstremnih vrednosti u prozoru za definisanje specifikacije, odlučuje se o tome da li se njihovo prisustvo potpuno odstranjuje ili se informacija koju oni nose uključuje u proces modeliranja.

#### 5.4.5 ARIMA model

Deo specifikacije je i odabir ARIMA modela. Korisnik može sam da definiše po svom mišljenju najreprezentativniji model, tj. parametre nesezonskog i sezonskog dela,  $(p,d,q)(bp, bd, bq)$ , koristeći se graficima autokorelace i parcijalne autokorelace funkcije, periodogramom i sl., ili iskustvima u dotadašnjim analizama ili karakteristikama serija na kojima radi. Sa druge strane, uvek se može ostaviti automatski odabir modela, kada program sam definiše model za koji se prepostavlja da najbolje prezentuje podatke. Na slici 5.20 su oba slučaja.



Slika 5.20: Odabir ARIMA modela

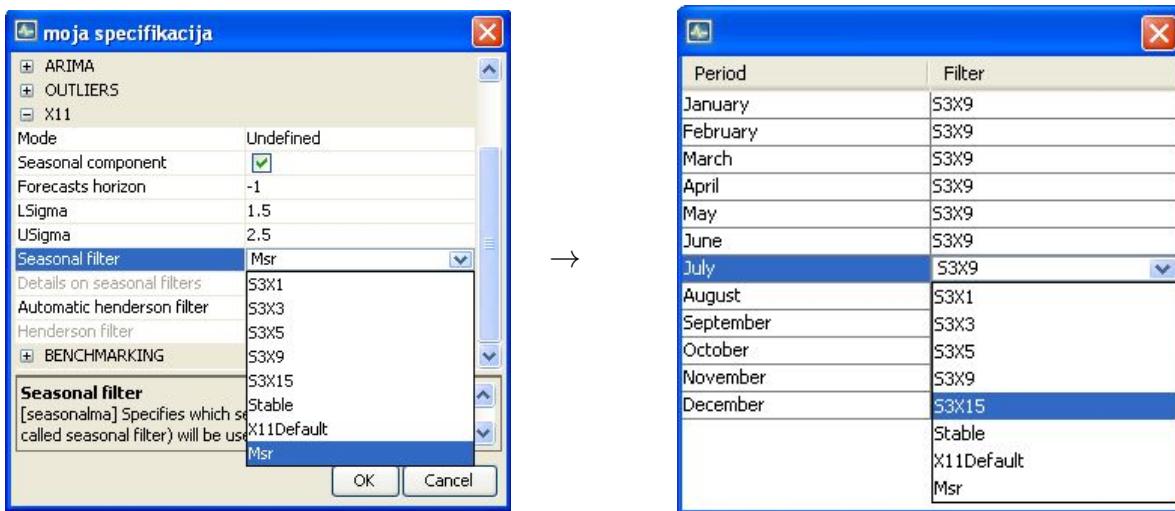
#### 5.4.6 Sezonski filter

Sezonski filter je ponderisani prosek pokretnih raspona fiksne dužine, koji se može koristiti za uklanjanje fiksног sezonskog obrasca neke vremenske serije. Mogućnost manualnog odabira sezonskog filtera ima samo *X-13ARIMA-SEATS* metod, ne i *TRAMO/SEATS*.

Uopšteno govoreći, automatski proces desezoniranja (sa *default* opcijama) uzima najadekvatnije pokretne proseke. Međutim, postoje slučajevi kada korisnik treba da specifira različite pokretne proseke u odnosu na one koje identificuje program. Na primer, iznenadni prekidi u sezonskom obrascu (recimo usled promena u metodologiji) mogu uticati negativno na automatski odabir najboljeg sezonskog filtera. U takvim slučajevima treba razmotriti korišćenje kratkih sezonskih filtera u odabranom mesecu ili kvartalu, jer oni dopuštaju da se sezonalnost brzo menja kroz vreme. Međutim, vrlo kratki sezonski filter ( $3 \times 1$ ) treba izbegavati, jer obično dovodi do velike revizije kad novi podaci postanu dostupni. Ovo se koristi samo za jedan mesec/kvartal i to kad nema razloga za praćenje brzih promena sezonalnosti.

U standardnoj proceduri, jedan sezonski filter koristi se za sve mesece/kvartale. Međutim, uzroci sezonalnosti mogu se promeniti u jednom mesecu, dok u drugima ostanu isti. Na primer, industrijska proizvodnja je uzrokovana i školskim raspustima, s obzirom na to da mnogi zaposleni imaju decu školskog uzrasta, pa je i više prekida rada tokom tog perioda. Kako školski raspusti nisu uvek u isto vreme širom Nemačke, a njihovi rasporedi variraju od godine do godine u različitim federalnim državama, efekat nije kompletno pokriven desezoniranjem. Takođe, ovi efekti nisu ni pokriveni kalendarskim prilagođavanjem, s obzirom na to da se školski raspusti računaju kao obični radni dani. Kako je većina školskih raspusta u Nemačkoj u julu i avgustu, to dovodi do većih varijansi u iregularnoj komponenti za ove mesece, u poređenju sa ostalima. Stoga, treba koristiti duži sezonski filter.

JDemetra+ nudi mogućnost podešavanja sezonskih filtera različitih dužina za svaki period. U specifikaciji, u delu *X-11*, nalazi se polje *Seasonal filter* gde se može odabrati željena dužina filtera. Polje *Details on seasonal filters* otvara novi prozor u kom se mogu predefinisati filteri za pojedine periode.



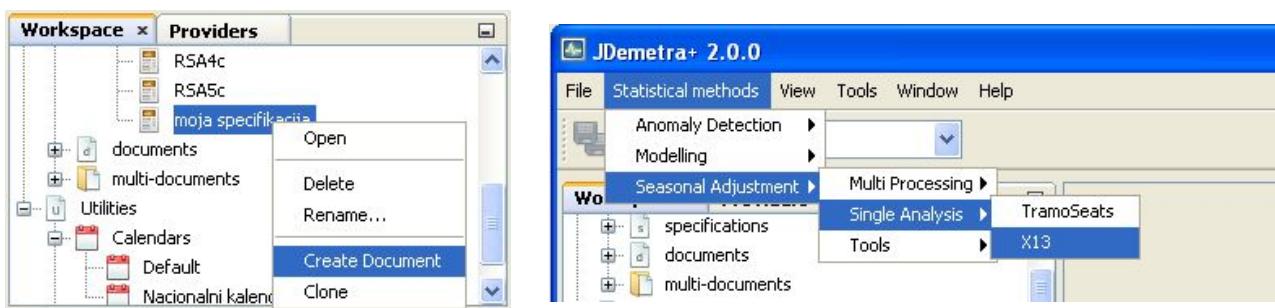
Slika 5.21: Odabir sezonskih filtera

U ovom odeljku specifikacije je i mogućnost odabira konkretnog Hendersonovog filtera, koji služi za izravnavanje linije trenda i uklanjanje svih kratkih iregularnih varijacija. On se koristi uvek posle otklanjanja sezonske komponente. Obično se prvo izvrši desezoniranje sa automatskim Hendersonovim filterom, pa se eventualno na pojedinačnim serijama vrše takve izmene, u cilju poboljšanja kvaliteta obrade.

## 5.5 Obrada jedne serije (engl. *Single processing*)

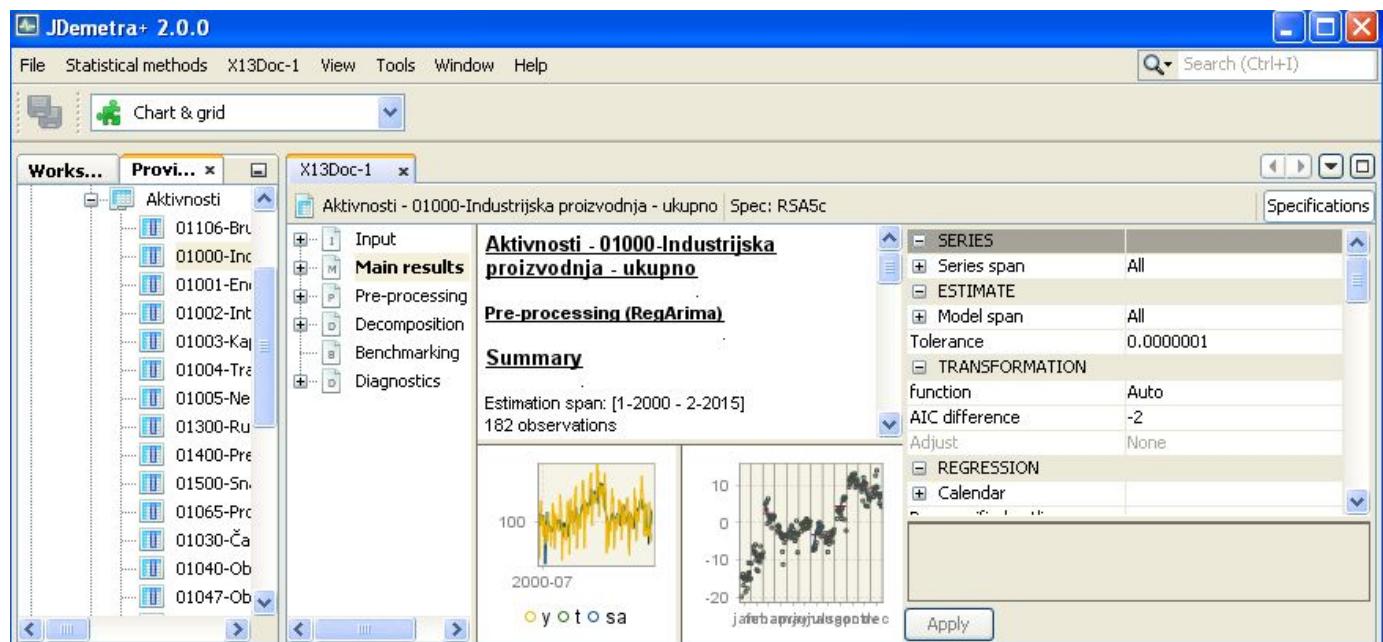
Kad je formirana željena specifikacija, sledeći korak je pokretanje obrade, tj. sam proces desezoniranja. Obradu jedne serije moguće je pokrenuti na dva načina:

- desnim klikom na 'moja specifikacija', pa odabirom opcije *Create Document*;
- preko linije menija: *Statistical methods* → *Seasonal Adjustment* → *Single Analysis* → *X13*.

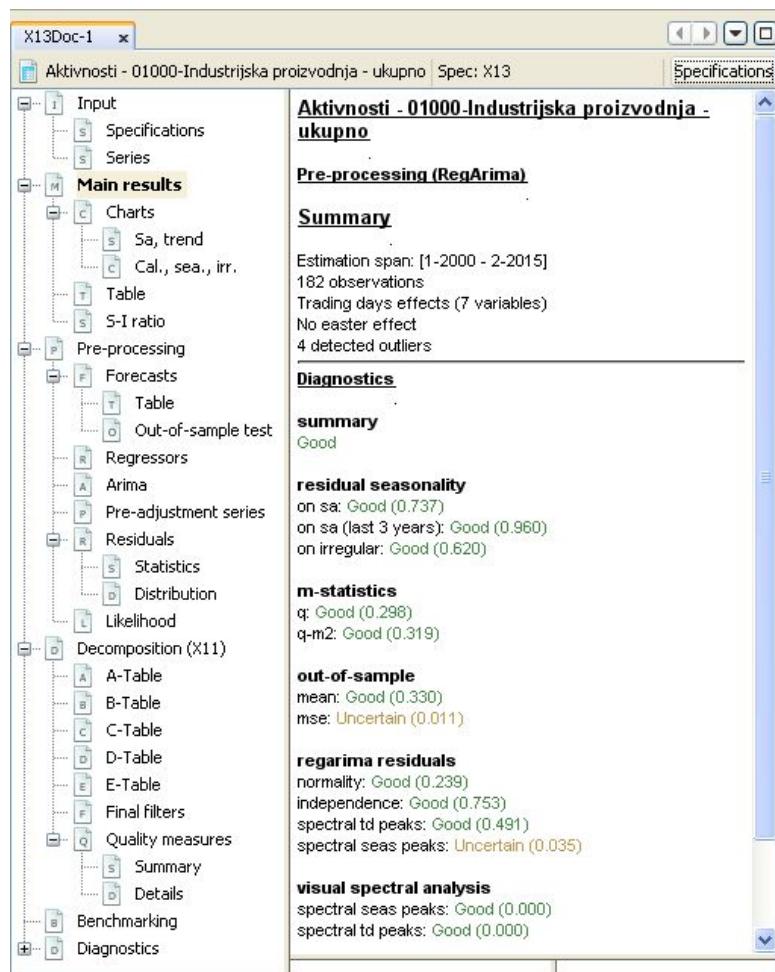


Slika 5.22: Pokretanje obrade jedne serije

Otvara se novi prozor u desnom delu glavnog prozora, gde se može prostim prevlačenjem miša (engl. *drag and drop*) preneti serija koja treba da se analizira. Proces desezoniranja se automatski pokreće, a prvi rezultati se odmah ispisuju. Korisnik ima mogućnost da menja specifikaciju, a time i celu obradu, aktiviranjem opcije *Specification* iz gornjeg desnog ugla. Svakom promenom specifikacije, klikom na dugme *Apply*, automatski se ponovo sprovodi desezoniranje. Na taj način korisnik može da upoređuje rezultate i odabere najreprezentativniji model.



Slika 5.23: Obrada jedne serije

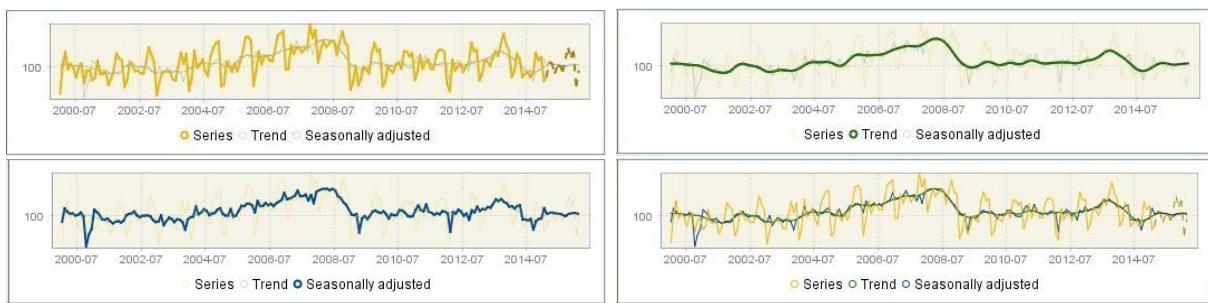


Slika 5.24: Main results

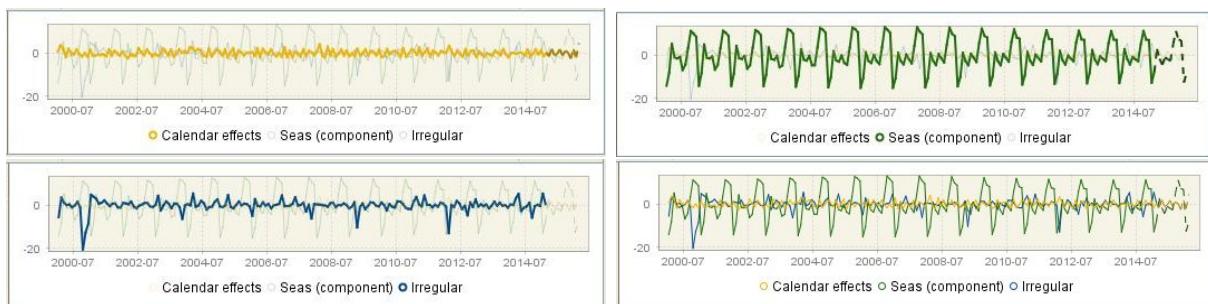
Na slici 5.24 opšti zaključak je dobar (engl. *Good*), što je sasvim prihvatljivo, ali ako korisnik želi, ima prostora da menjanjem specifikacije dotera i niže grane dijagnostike i time usavrši kvalitet desezoniranja.

JDemetra+ daje iscrpne i vrlo detaljne informacije o seriji, kao i obimne testove za proveru kvaliteta procesa desezoniranja. Svi rezultati predstavljeni su u pet glavnih celina: Ulaz (engl. *Input*), Glavni rezultati (engl. *Main results*), Preprocesiranje (engl. *Pre-processing*), Dekompozicija (X11) (engl. *Decomposition (X11)*) i Dijagnostika (engl. *Diagnostics*). U *Input*-u je odabrana specifikacija i tabela sa ulaznim podacima. U *main results*-u nalaze se kratki opisi modela koji se koristi i kvaliteta desezone. Najpre je dat podatak o vremenskoj distanci serije koja se ocenjuje, broj observacija, kao i informacija o eventualnoj logaritamskoj transformaciji, prisustvu efekta kalendara i detektovanim autolajerima. Zatim je predstavljen sumarni uvid u kvalitet analize. Mogući pokazatelji su *Good*, *Uncertain* i *Severe*. Prvi ne iziskuje nikakve izmene, kod drugog su poželjne, ali ne i obavezne, a kod trećeg su neophodne, sudeći po činjenici da program statisistički ne prihvata kvalitet odabranih parametara.

U ovom odeljku nalaze se i grafički i tabelarni prikazi serije koja je analizirana. Korisnik ima mogućnost da odvojeno vidi svaku od komponenti, kao i sve njih na istom grafiku.

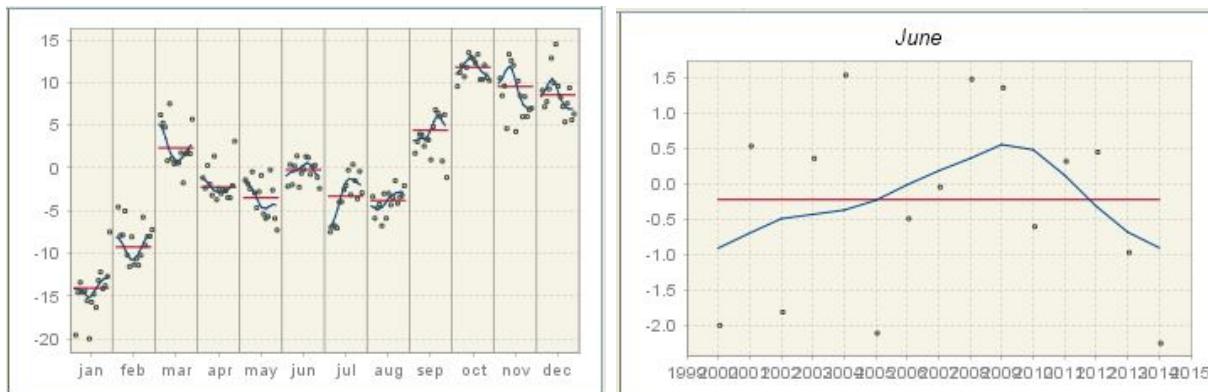


Slika 5.25: Main results → Charts → Sa, trend



Slika 5.26: Main results → Charts → Cal., sea., irr.

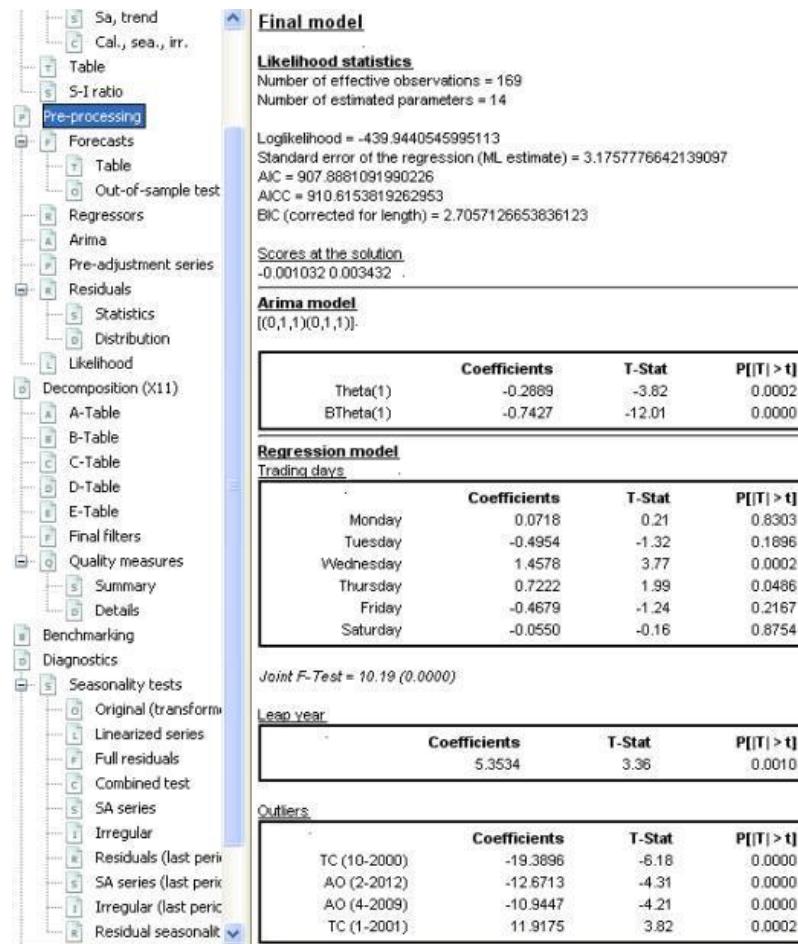
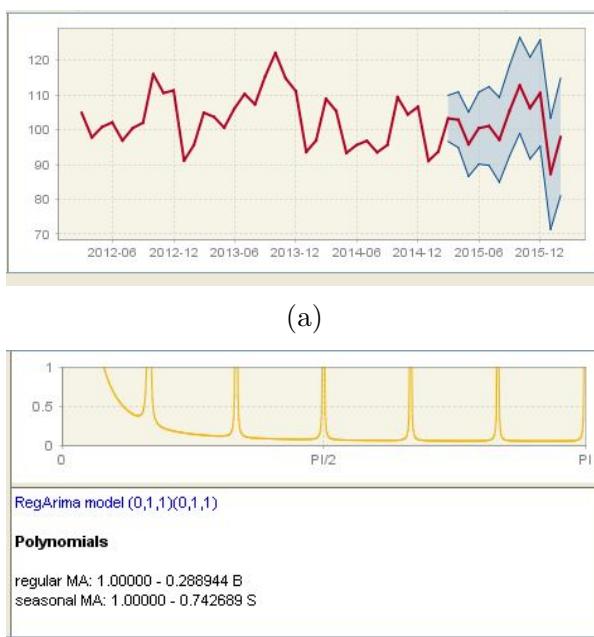
Jedan od najznačajnijih deskriptivnih pregleda dekompozicije je i *S-I ratio chart*, koji ilustruje nestabilne ili pokretne sezonske faktore, ako postoje. Koristan je za otkrivanje sezonskih pauza i perioda sa visokom varijabilnošću. Sezonske pauze mogu izazvati probleme sa pokretnim proseccima, identifikovanjem *trading* dana i efekta Uskrsa, kao i odgovarajućim ARIMA modelom. Takve serije su komplikovane za desezoniranje i uglavnom bi ih trebalo podeliti na dva dela, pa ih odvojeno analizirati. Dvostruki klik na specifični period otvara detaljniji pregled sezonske staze jednog meseca ili kvartala (slika 5.27). Visok stepen promena u observacijama reflektuje prisustvo pokretne sezonalnosti, što znači da strme promene izazivaju sezonske prekide. Tačkice su S-I odnosi (engl. *S-I ratios*), a izračunavaju se kao odnos originalne serije i trend komponente, tj. to su ocene detrendovane serije (kako je  $sa = \frac{y}{s}$  i  $sa = i \cdot t \Rightarrow \frac{y}{s} = i \cdot t \Rightarrow si = \frac{y}{t}$ ). Plave linije pokazuju finalne sezonske faktore, a crvene prosek za svaki mesec.



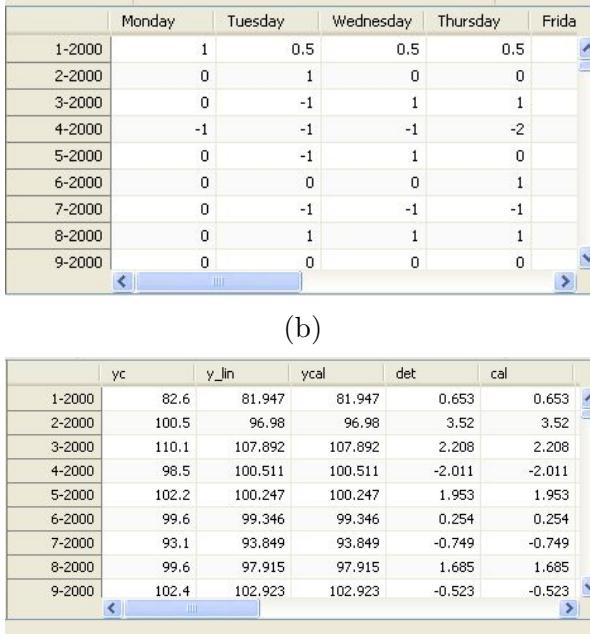
Slika 5.27: S-I ratio chart

U *Pre-processing*-u se nalazi ocena modela po raznim kriterijumima, među kojima su i Akaikeov i Bajesov informacioni kriterijum. Takođe su prikazani rezultati detektovanja autlajera (TC, LS i AO), statistička značajnost *trading* dana i prestupne godine, kao i rezultati *JointF* testa, koji utvrđuje da li je bolje koristiti *trading* ili *working* dane. Ovde je naznačen i Arima model, koji je po programu najrepresentativniji za učitane podatke.

U *Forecasts*-u su date prognoze (grafički i tabelarno) i njihovi intervali poverenja (sl. 5.29(a)). Klikom na *Regressors* dobija se tablica determinističkih regresora, koji su korišćeni u RegArima delu (*trading* dani, prestupna godina, autlajeri, Uskrs, promenljive koje je korsnik definisao itd.) (sl. 5.29(b)), a *Arima* daje spektralni grafik sa pikovima na frekvencama za mesečne serije i ocnjene koeficijente parametara RegArima modela (sl. 5.29(c)). Opcija *Pre-adjustment series* tabelarno predstavlja serije ocnjene takođe u RegArima delu, od kojih su najbitnije kalendarska komponenta (*cal*) i serija korigovana za efekat kalendara, tj. kalendarski prilagođena serija (*ycal*) (sl. 5.29(d)).

Slika 5.28: *Pre-processing*

(a)

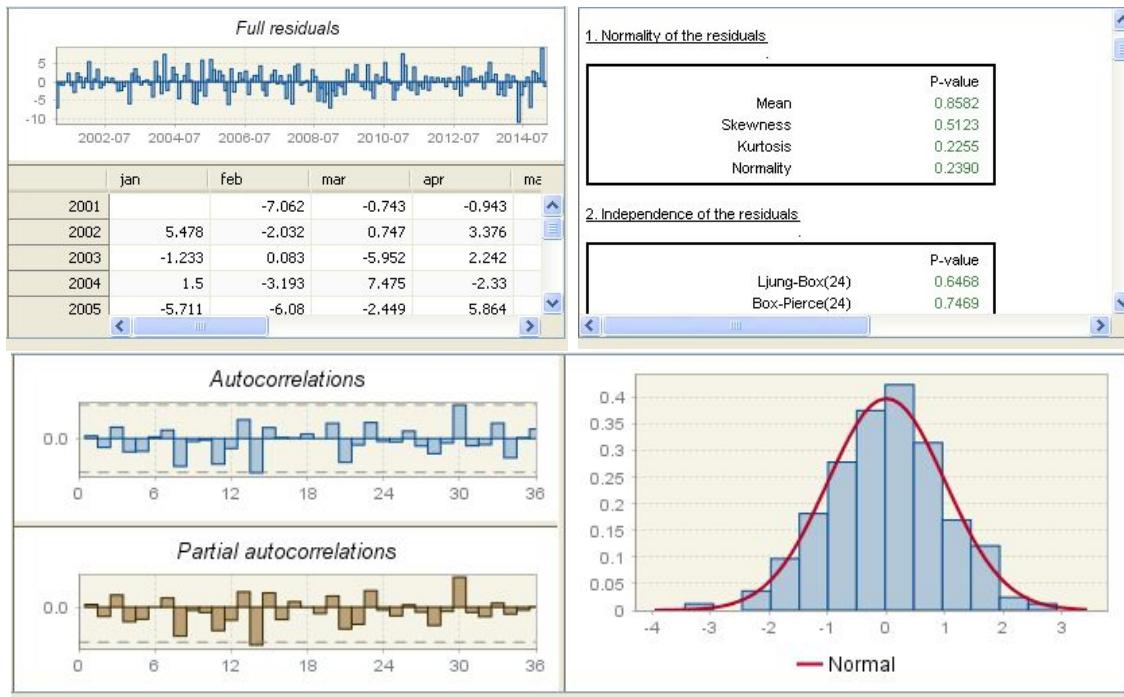


(d)

Slika 5.29: *Forecasts, Regressors, Arima, Pre-adjustments series*

Reziduali su takođe u ovom odeljku, kao vrlo korisni indikatori za proveru kvaliteta desezoniranja. Svi su izlistani tabelarno i grafički. Trebalo bi da budu nezavisni i slučajni

i normalno raspodeljeni. Prosek bi trebalo da je nula, jer ne bi trebalo da nose neku informaciju, poput sezonalnosti. Dostupni su i statistički testovi koji ispituju ove osobine reziduala. Proverava se normalnost, nezavisnost, slučajnost i linearost reziduala, a p-vrednosti testova su obojene zelenom (prihvata se nulta hipoteza sa pragom značajnosti od 5%), žutom (prihvata se nulta hipoteza sa pragom značajnosti od 1%) ili crvenom bojom (ne prihvata se nulta hipoteza). Takođe, korisnik može da vidi i grafike autokorelacione i parcijalne autokorelacione funkcije, kao i histogram raspodele reziduala. Slika 5.30 pokazuje ove opcije.



Slika 5.30: *Residuals*

U *Decomposition (X11)* nalaze se različite tablice, čije se vrednosti dobijaju računima *X11* metode:

- *A-Tablica* - preliminarna ocena ekstremnih vrednosti i efekta kalendarja;
- *B-Tablica* - preliminarna ocena komponenata vremenske serije;
- *C-Tablica* - finalna ocena ekstremnih vrednosti i efekta kalendarja;
- *D-Tablica* - finalna ocena različitih komponenti;
- *E-Tablica* - komponente modifikovane velikim ekstremnim vrednostima;

Najkorisnije među njima su kolona *A1* - originalna serija (*y*), *D11* - finalna desezonirana serija (*sa*), *D12* - trend komponenta (*t*), *D10* - sezonska komponenta (*s*), *C16B* - efekat kalendara (*cal*) i *D13* - iregularna komponenta (*i*).

U *Final filters*-u je definisan sezonski i trend (Hendersonov) filter korišćen za model.

### Final filters

Seasonal filter: 3x5  
Trend filter: 13 terms Henderson moving average

Slika 5.31: *Final filters*

*Quality measures* odeljak daje vrednosti  $M$ -statistika, koje treba da su manje od jedan za dobre rezultate.  $Q$  i  $Q$  bez  $M$  su kompozitni indikatori, izračunati na osnovu  $M$  statistika. Oni, takođe, treba da su manji od jedan da bi kvalitet desezoniranja bio zadovoljavajući.

<b>Monitoring and Quality Assessment Statistics</b>		
M-1	0.152	The relative contribution of the irregular over three months span
M-2	0.130	The relative contribution of the irregular component to the stationary portion of the variable
M-3	0.394	The amount of period to period change in the irregular component as compared to the amount of stable seasonality
M-4	0.617	The amount of autocorrelation in the irregular as described by the average duration of runs
M-5	0.586	The number of periods it takes the change in the trend- cycle to surpass the amount of stable seasonality
M-6	0.111	The amount of year to year change in the irregular as compared to the amount of year to year change in the seasonal component
M-7	0.191	The amount of moving seasonality present relative to the amount of stable seasonality
M-8	0.416	The size of the fluctuations in the seasonal component throughout the whole series
M-9	0.159	The average linear movement in the seasonal component throughout the whole series
M-10	0.384	The size of the fluctuations in the seasonal component in the recent years
M-11	0.360	The average linear movement in the seasonal component in the recent years
<b>Q</b>	<b>0.298</b>	
<b>Q-m2</b>	<b>0.319</b>	

Slika 5.32: *Quality measures*

Dijagnostika (engl. *Diagnostics*) prezentuje detaljne informacije o proceduri desezoniranja. Podeljena je na pet glavnih delova: testove za prisustvo sezonalnosti, spektralne grafike, klizne raspone, istoriju revizije i analizu stabilnosti modela. Prvi deo služi za utvrđivanje postojanja sezonskog obrasca, tj. potrebe za desezoniranjem. Rezultati različitih testova nisu uvek usaglašeni, pa se kao najkompleksniji i najmerodavniji za procenu uzima kombinovani test.

**Non parametric tests for stable seasonality**

**Friedman test**

Friedman statistic = 151.2974  
Distribution: Chi2(11)  
P-Value: 0.0000

Stable seasonality present at the 1 per cent level

**Kruskall-Wallis test**

Kruskall-Wallis statistic = 165.15565213475054  
Distribution: Chi2(11)  
P-Value: 0.0000

Stable seasonality present at the 1 per cent level

**Test for the presence of seasonality assuming stability**

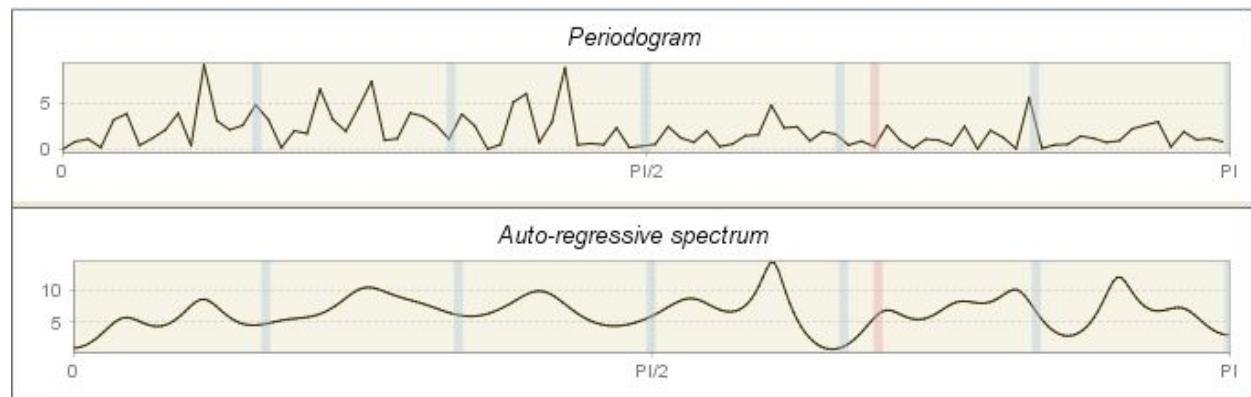
	Sum of squares	Degrees of freedom	Mean square
Between months	9697.191263976441	11.0	881.5628421796765
Residual	865.8100554707617	170.0	5.093000326298598
Total	10563.001319447203	181.0	58.359123311863

Value: 173.09302684069596  
Distribution: F with 11 degrees of freedom in the nominator and 170 degrees of freedom in the denominator  
PValue: 0.0000  
Seasonality present at the 1 per cent level

**Evolutive seasonality test**

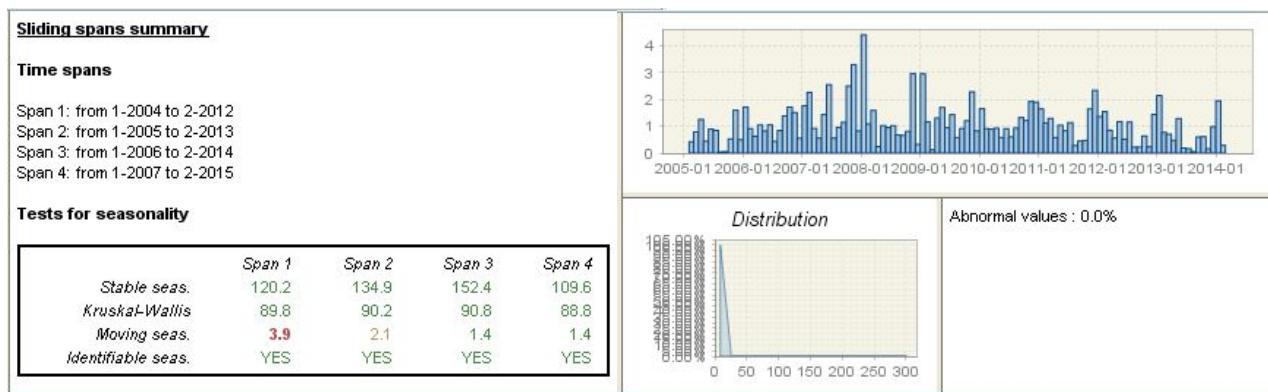
Slika 5.33: *Diagnostics → Seasonality tests*

Spektralna analiza obuhvata spektralne grafove, tj. periodogramme i autoregresione spektre reziduala, iregularne komponente i desezonirane serije. Njima se proverava prisustvo pikova na sezonskim frekvencama (sive linije) i frekvencama *trading* dana (crvena linija). Pikovi na sezonskim frekvencama desezonirane serije znače da filteri korišćeni u dekompoziciji nisu dobro prilagođeni seriji ili njenom većem delu. Pikovi na frekvencama *trading* dana mogu da ukazuju da regresione promenljive modela ne pristaju dobro seriji ili da je efekti kalendara toliko menjaju da ne mogu biti obuhvaćeni samo fiksnim regresionim efektima primjenjenim na ceo tok serije. Ako ima zaostale sezonalnosti ili efekta *trading* dana treba ponovo razmotriti specifikaciju, regresione promenljive ili dužinu serije koja se modelira (engl. *model span*). Svi ovi grafici mogu se kopirati i eksportovati.



Slika 5.34: Diagnostics → Spectral analysis

Klizni rasponi (engl. *Sliding spans*) analiziraju stabilnost desezonirane serije, koja se ogleda u tome da dodavanje ili oduzimanje vrednosti na početku ili kraju serije ne utiče mnogo na *sa* seriju. Ovim rasponima mogu se otkriti značajne promene u serijama, poput sezonskih prekida, velikog broja autolajera i brzih pokretnih sezonalnosti. Zavisno od dužine originalne serije, klizni rasponi uključuju dva, tri, ili četiri preklapajuća raspona, a za svaki od njih program posebno radi testove sezonalnosti. Više detalja dostupno je za sezonsku komponentu, efekat *trading* dana i desezoniranu seriju. Mogu se videti statistike kliznih raspona za svaku od njih, raspodela raspona i tablica sa nestabilnim faktorima. Ako ove statistike pokazuju veliki broj nestabilnih ocena, treba razmisiliti o promeni specifikacije modela.



Slika 5.35: Sliding spans

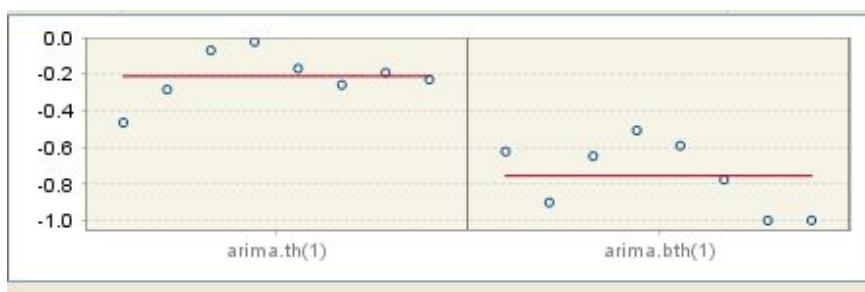
U *Revisions history* predstavljeni su grafici desezonirane serije i trend komponente. Svaki plavi kružić prikazuje početno prilagođavanje kad je ova tačka poslednja observacija. Crvena linija su finalni rezultati. Analiza polazi od ocene modela za ceo vremenski raspon. Nadalje,

vremenski raspon se progresivno skraćuje, a dekompozicija ponovo ocenjuje. Za svaki period dobijena je serija uzastopnih procena. Po *default*-u, ponovo se procenjuju samo parametri modela. Međutim, program omogućava potpunu ponovnu procenu i reidentifikaciju autlajera ako se opcija promeni. Klikom na tačkicu na grafikonu, pojavljuje se *pop-up* prozor, koji pokazuje uspešne ocene te vremenske tačke za različite vremenske raspone. Na slici 5.36 potvrđuje se da revizija nije značajna u poslednjih godinu dana. U ovom delu nalazi se još i tablica koja predstavlja razlike između prvih i poslednjih ocena za poslednje četiri godine. Ako je dekompozicija aditivna, prikazana je absolutna revizija, a inače se koriste relativne revizije. Najveće razlike su u crvenoj boji i one su veće od dvostrukog korena srednje kvadratne greške absolutne ili relativne revizije. Obično datu kritičnu vrednost prelaze autlajeri i oni su crveni.



Slika 5.36: *Revisions history*

Poslednji deo dijagnostike je analiza stabilnosti modela, gde se računaju ARIMA parametri i koeficijenti regresionih promenljivih za različite periode. Isto i za efekat *trading* dana i Uskrsa. Rezultati su dati vizuelno. Za seriju datu na slici 5.37 uključeni su vremenski rasponi 2000 – 2008, 2001 – 2009, ..., 2006 – 2014, 2007 – 2015 - ukupno osam tačaka. Od tога koliko су one stišnjene oko crvene linije zavisi kvalitet odabranog modela.

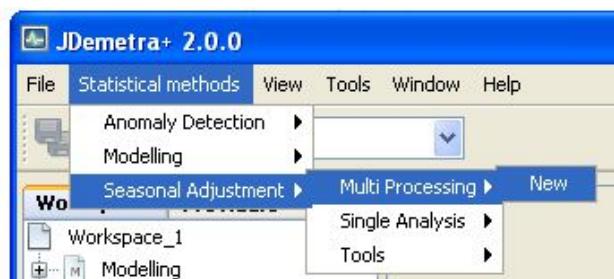


Slika 5.37: *Model stability → Arima*

Još jedna od opcija, do sada nespomenutih, je *Benchmarking*. To je procedura koja obezbeđuje prisilnu konzistentnost tokom godine između desezoniranih i nedesezoniranih podataka. Obično se ne preporučuje, jer dovodi do pristrasnosti u procesu desezoniranja i time ugrožava kvalitet analize. Ni serije Republike Srbije ne uključuju *benchmarking* metod, jer on izobliči kretanje serije, pa u daljem tekstu njegovo učešće u procesu desezoniranja neće biti razmatrano.

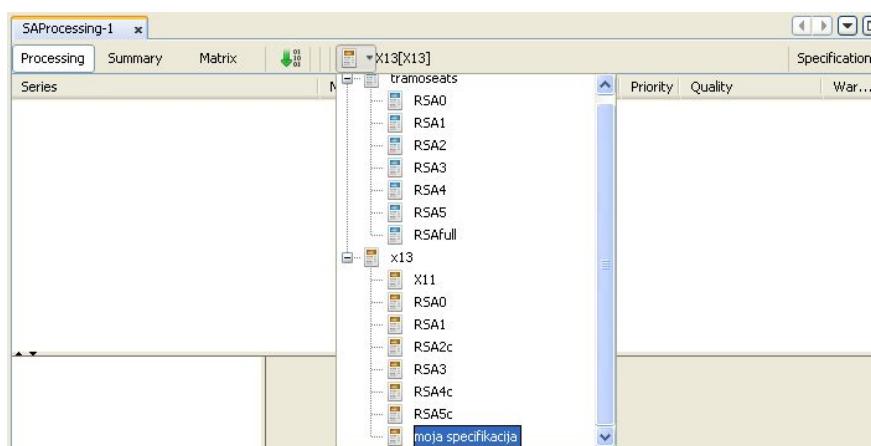
## 5.6 Obrada grupe serija (engl. *Multi processing*)

U slučaju kada odjednom treba desezonirati veću grupu serija, srodnih ili ne, korisno je upotrebiti *multi-processing*. Pri tom se sve serije mogu desezonirati korišćenjem opcija iste specifikacije ili se podvojeno, za pojedinačne grupe serija ili samo jednu seriju, kreiraju specijalne specifikacije (kako je pokazano u poglavlju 5.4).



Slika 5.38: Pokretanje procesa obrade grupe serija

U glavnom meniju, staza prizana na slici 5.38, otvara prozor u kom će se vršiti obrada grupe serija. Po *default*-u se unapred definisana specifiakcija *TS/RSAfull* koristi za desezoniranje, ali, naravno, korisnik to može samoinicijativno promeniti (slika 5.39).



Slika 5.39: Promena specifikacije pri grupnoj obradi serija

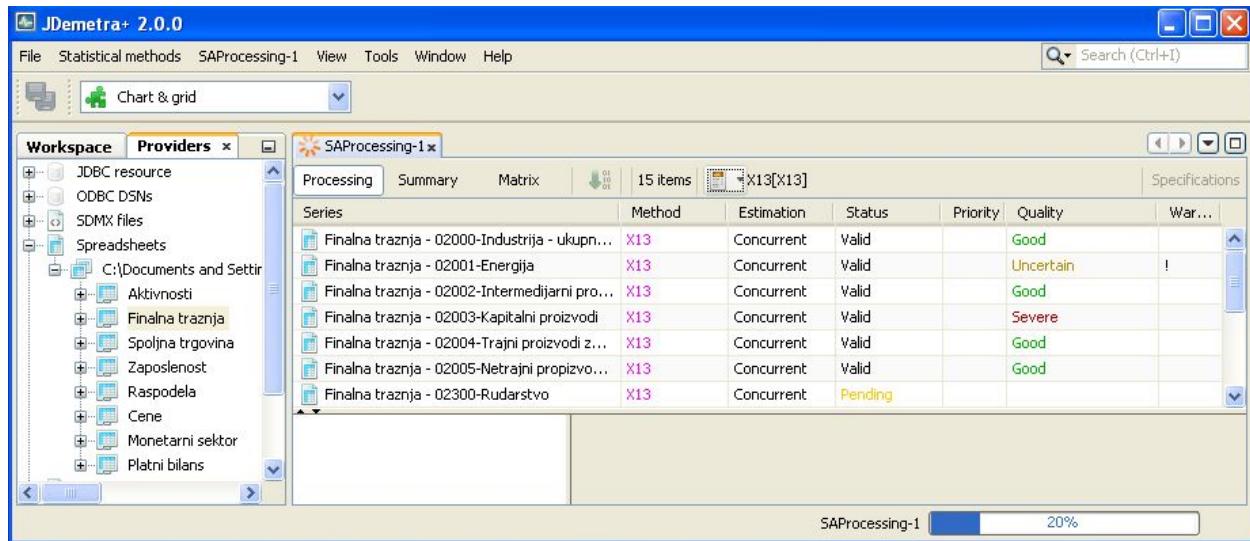
Prostim prevlačenjem serija iz *Providers* panela u *SAPProcessing* prozor vrši se odabir serija na kojima će se vršiti obrada. Moguće ih je individualno prenositi ili ceo ulazni fajl preneti odjednom. Takođe, dozvoljeno je istu seriju preneti više puta, što je korisno ako poredimo dva različita tipa obrade. Serije vidljive u prozoru *SAPProcessing* (slika 5.40) još uvek nisu desezonirane (njihov status je *Unprocessed*).

Series	Method	Estimation	Status	Priority	Quality	War...
Finalna traznja - 02000-Industrija - ukup...	X13	Concurrent	Unprocessed			
Finalna traznja - 02001-Energija	X13	Concurrent	Unprocessed			
Finalna traznja - 02002-Intermedijni pro...	X13	Concurrent	Unprocessed			
Finalna traznja - 02003-Kapitalni proizvodi	X13	Concurrent	Unprocessed			
Finalna traznja - 02004-Trajni proizvodi z...	X13	Concurrent	Unprocessed			
Finalna traznja - 02005-Netrajni propizvo...	X13	Concurrent	Unprocessed			
Finalna traznja - 02300-Rudarstvo	X13	Concurrent	Unprocessed			

Slika 5.40: Odabir serija za desezoniranje

Obrada će biti izvršena korišćenjem metode *X-13* i to specifikacije *moja specifikacija*. Kolona *Estimation* za novodefinisani *SAPProcessing* prozor uvek je postavljena na *Concurrent*, što znači da prethodni rezultati za ove vremenske serije neće biti uzeti u obzir kad se pokrene proces desezoniranja (s obzirom na to da i ne postoje). Ako se ovaj prozor sačuva i ponovo pokrene, korisniku je dozvoljeno da odabere da li će se i kako prethodni rezultati računati u tekućoj obradi, o čemu će kasnije biti više reči. Kolone *Priority*, *Quality* i *Warnings* su prazne dok se proces desezoniranja ne pokrene.

Zelenom strelicom, , pokreće se proces desezoniranja.

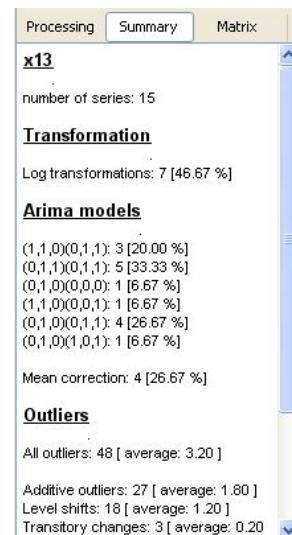


Slika 5.41: Desezoniranje grupe serija

Statusi se ažuriraju, a neke informacije o kvalitetu desezoniranja i eventualni problemi ili upozorenja se pojavljuju u koloni *Warnings* (oznaka je znak užvika). Obično se upozorenja javljaju kad su serije prekratke ili kada diferencirane serije nemaju sezonske vrhove. Klikom na pojedinačnu seriju u *SAPProcessing* prozoru, prikazuju se detaljni rezultati obrade u panelima ispod liste serija, identični onima iz prethodnog poglavlja za pojedinačnu obradu. Svi rezultati obojeni su zeleno (*Good*), žuto (*Uncertain*) ili crveno (*Severe*), pa se u zavisnosti od toga odlučuje da li i kakve izmene na pojedinačnim specifikacijama, dostupnim u gornjem desnom uglu *SAPProcessing* prozora, treba vršiti.

*Summary* (slika 5.42) daje opšte rezultate grupne obrade serija. *x13* metodom desezonirano je ukupno 15 serija, od kojih je 7 logaritamski transformisano, a među kojima je najzastupljeniji vazduhoplovni model. Ukupno je 48 autlajera, što je prosečno tri po seriji, a to ulazi u domet očekivanog. Efekat kalendara uočen je na svim serijama, jer je prethodno ubačen nacionalni kalendar, koji se koristi u svakoj obradi.

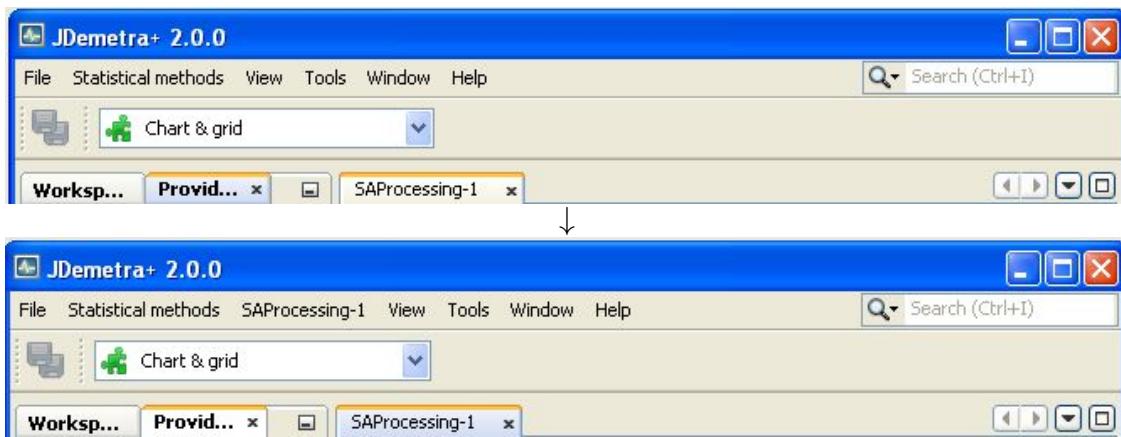
*Matrix* sadrži sumarne informacije za svaku seriju o njenim statističkim karakteristikama modela, strukturi autlajera, efekta kalendara i testova značajnosti (koeficijenti, regresori ili p-vrednosti). Svaka od ovih matrica može se kopirati ili eksportovati u bilo koji drugi dokument.



Slika 5.42: Opšti rezultati

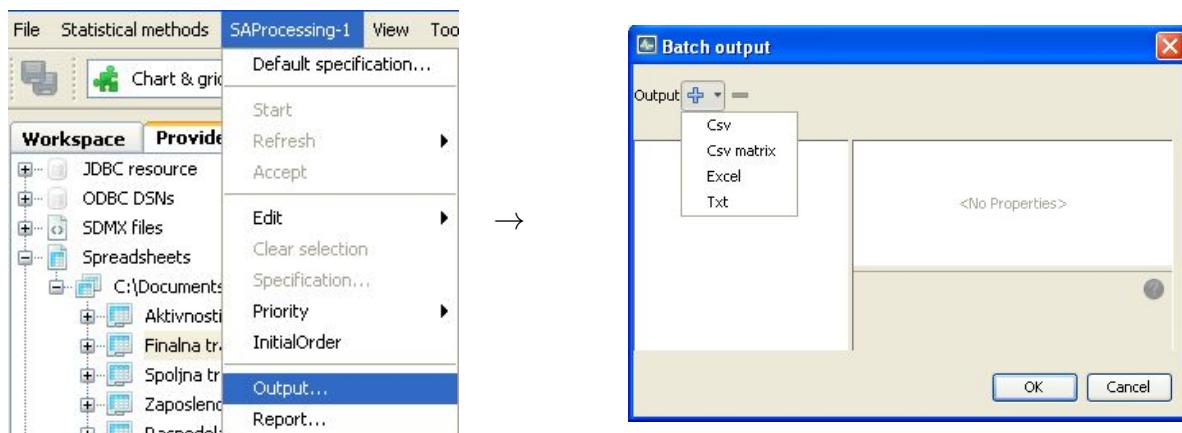
## 5.7 Kreiranje izlaznog fajla sa rezultatima (engl. *Output*)

Svi rezultati desezonirane grupe serija mogu se eksportovati u neki spoljni fajl, i to svi odjednom. Klikom na prozor *SAProcessing* pojavljuje se novi meni u liniji menija, između *Statistical methods* i *View*:



**Slika 5.43:** Promena linije menija usled aktiviranja SAProcessing prozora

Korisnik može da odabere ekstenziju fajla u koji želi da eksportuje rezultate. Dozvoljeno je kreiranje tekstualnog (.txt), Excel (.xlsx), csv (*Comma separated values*) ili matričnog csv fajla. Staza SAProcessing → Output otvara Batch output prozor, gde je upravo taj odabir prvi korak u eksportovanju.



**Slika 5.44:** Eksportovanje rezultata

S obzirom na to da je uobičajeno izlaze praviti u Excel-u, taj će proces i biti pojašnjen. U desnom panelu *Batch output* prozora otvara se *Properties* za definisanje formata i oblika izlaza (slika 5.45).

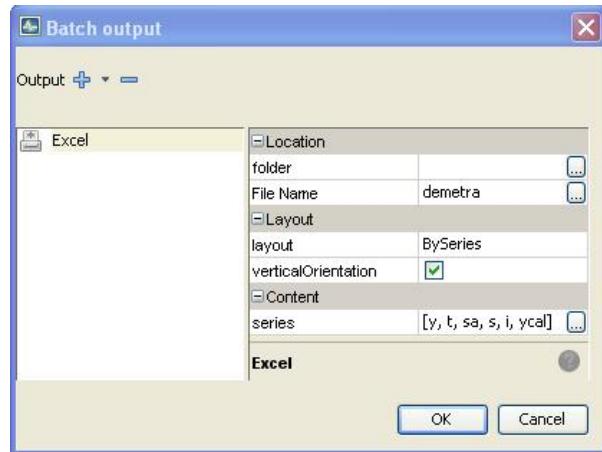
Prvo polje (*folder*) služi za definisanje lokacije na kojoj će biti smešten kreirani .xlsx fajl sa izlaznim rezultatima. U drugom polju (*File Name*) zadaje se ime fajla. Odeljak *Layout* nudi tri opcije:

- *BySeries* - svi rezultati za datu vremensku seriju smeštaju se u jedan *excel sheet*, tj. svaka vremenska serija ima svoj *excel sheet*;
- *ByComponent* - rezultati su grupisani po komponentama, tj. svaka komponenta je u različitom *excel sheet*-u;

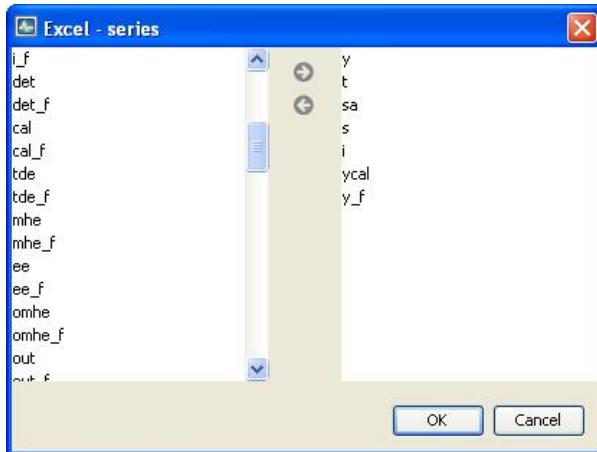
- *OneSheet* - svi rezultati svih serija su u istom *excel sheet-u*.

Po *default*-u su serije u izlaznom Excel fajlu vertikalno orijentisane. Ukoliko korisnik želi to da promeni, dečekiranjem opcije *verticalOrientation* i to je omogućeno.

Naposletku, polje *Content* → *series* služi za odabir željenih komponenti i serija u izlazu. Najčešće su potrebne originalna serija (*y*), trend (*t*), desezonirana serija (*sa*), sezonska komponenta (*s*), iregularna komponenta (*i*), kalendarски prilagođena serija (*ycal*) i prognoza originalne serije (*y-f*) (slika 5.46). Strelicama se vrši odabir, a pored ovih, dostupne su i prognoze svih komponenti, autlajeri, efekti kalendarja, regresori i determinističke komponente.

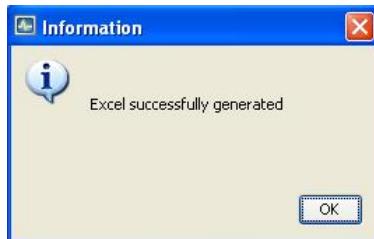


Slika 5.45: Opcije za kreiranje izlaza



Slika 5.46: Komponente za izlaz

Jednom kad korisnik postane zadovoljan definisanim opcijama i formatom izlaza, potvrđivanjem akcije dobija se obaveštenje sledećeg sadržaja:



Slika 5.47: Informacija o generisanom izlazu

Na zadatoj lokaciji, pod zadatim imenom, kreiran je željeni Excel fajl. Neke od mogućih varijanti izlaza date su na sledećoj slici:

	A	B	C	D	E
1		Finalna tražnja - 02000-Industrija - ukupno			
2	<i>y</i>	<i>t</i>	<i>sa</i>	<i>s</i>	
3	2000-01-01	105	108.0956	104.8766	1.0011
4	2000-02-01	108.8	108.0892	108.2308	1.0052
5	2000-03-01	106.9	108.0383	107.9211	0.9905
6	2000-04-01	106	107.956	107.6757	0.9844
7	2000-05-01	106.9	107.8114	108.1184	0.9887
8	2000-06-01	107.6	107.4943	107.8843	0.9973
9	2000-07-01	105.6	106.9359	107.0851	0.9861
10	2000-08-01	102.8	106.2102	106.4309	0.965885

(a) *BySeries*

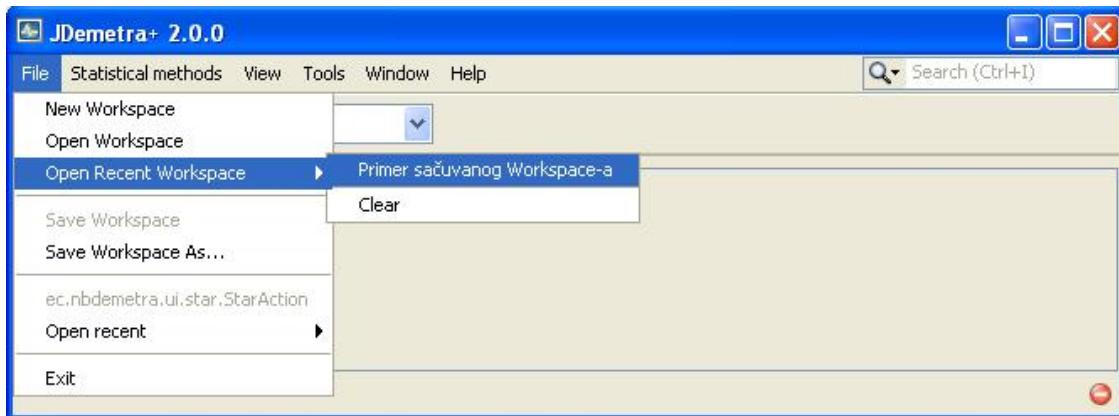
(b) *ByComponent*

(c) *OneSheet*

Slika 5.48: Primeri izlaznih Excel fajlova

## 5.8 Revizija procesa desezoniranja (engl. *Refresh options*)

Sačuvani rezultati procesa desezoniranja mogu se osvežiti kada su dostupni novi ili izmenjeni podaci. U tom slučaju se po pokretanju alata ponovo poziva ranije napravljeni *Workspace* sa lokacije na kojoj je prethodno sačuvan (opcije *Open Workspace* ili *Open Recent Workspace*).



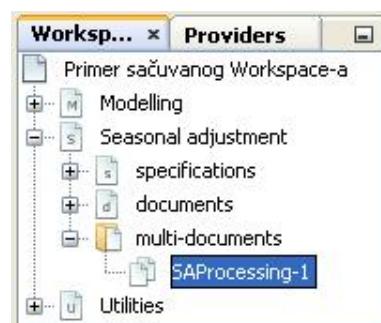
Slika 5.49: Otvaranje postojećeg *Workspace-a*

Potvrda da je stari *Workspace* otvoren je to što panel dobija njegovo ime.



Slika 5.50: Promena imena *Workspace* prozora nakon pozivanja postojećeg

Staza *Workspace Window* → *Seasonal adjustment* → *multi-documents* → *SAProcessing-1* otvara staru obradu. Excel fajl iz koga su podaci učitani u JDemetru+ ne sme menjati ni sada lokaciju, niti svoje ime, već se samo mogu vršiti izmene unutar njega (dodati/oduzeti/menjati podaci). U suprotnom će revizija biti onemogućena i pojaviće se greška u čitanju starog *Processing-a*.



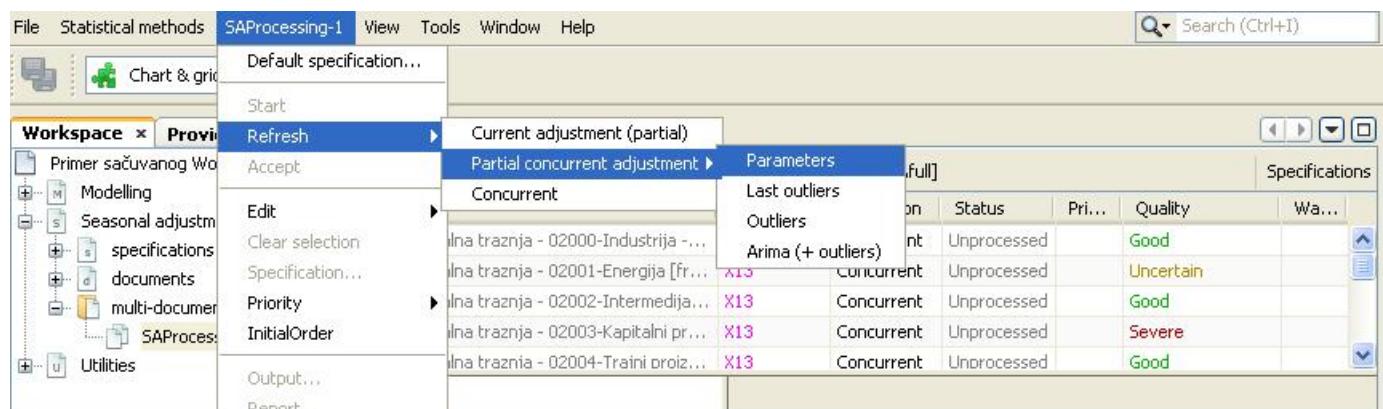
Slika 5.51: Stari *Processing*

JDemetra+ nudi nekoliko opcija za osvežavanje rezultata (slika 5.52):

- *Current adjustment (partial)* - ponovo se identifikuju i ocenjuju model, filteri, autlajeri i regresioni parametri u skladu sa razmatranim periodom, a sezonski i efekti kalendara koji se koriste za nove podatke su oni ocenjeni u prethodnom periodu i prognozirani za sledeći period ispitivanja; drugčije rečeno, ponovo se ocenjuju koeficijenti regresionog dela RegARIMA modela (autlajeri i efekti kalendara) - prethodno identifikovani autlajeri i efekti kalendara ostaju u modelu, ne dodaju se novi autlajeri i ne menja se tip

postojećih, već se samo ocenjuju nove vrednosti koeficijenata za ove efekte, ali bez ikakvih intervencija sa ARIMA delom modela;

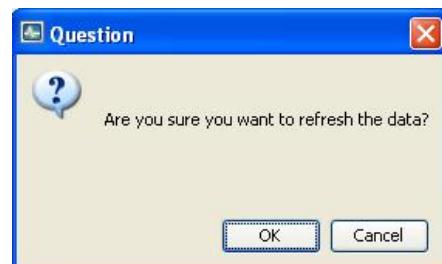
- *Partial concurrent adjustment → Parameters* - ponovo se ocenjuju parametri RegARIMA modela, tj. autlajeri, efekti kalendarja i parametri ARIMA modela; prethodno identifikovani ostaju nepromjenjeni;
- *Partial concurrent adjustment → Last outliers* - ponovo se ocenjuju koeficijenti regresionog dela modela (autlajeri i efekti kalendarja), a prethodno identifikovani ostaju u modelu; softver testira autlajere u poslednjoj godini, identificiše ih i ocenjuje; ocenjuju se nove vrednosti koeficijenata za efekte kalendarja, stare autlajere i ARIMA deo modela, takođe;
- *Partial concurrent adjustment → Outliers* - prethodno identifikovani efekti kalendarja ostaju u modelu, a njihove vrednosti parametara se ponovo ocenjuju; ARIMA deo modela je isti kao pre, a samo se koeficijenti ponovo ocenjuju; prethodno detektovani autlajeri se ne uzimaju u obzir, već se svi autlajeri identificišu i ocenjuju ponovo;
- *Partial concurrent adjustment → Arima (+outliers)* - prethodno identifikovani efekti kalendarja ostaju u modelu, a vrednosti parametara se ponovo ocenjuju; ARIMA deo modela se ponovo identificiše, što znači da može da se promeni; prethodno detektovani autlajeri se ne uzimaju u obzir, već se svi ponovo identificišu i ocenjuju;
- *Concurrent adjustment* - ponovo se identificišu model, filteri, autlajeri i regresioni parametri, a odgovarajući faktori se ponovo ocenjuju svaki put kada je nova observacija dostupna; drugim rečima, softver identificiše potpuno novi model, ne uzimajući u obzir prethodne rezultate.



Slika 5.52: Osvežavanje rezultata

Po odabranoj opciji, softver upozorava na predstojeće izmene (slika 5.53).

Za osvežavanje rezultata serija Republike Srbije mesečno se vrši *Partial concurrent adjustment → Parameters* revizija, jer se održava konzistentnost tokom godine, a jednom godišnje (obično na početku), kada se vrši i ažuriranje kalendarja, bira se novi model i vrši se *Concurrent adjustment* revizija. Naravno, ako se nakon osvežavanja utvrди da su neki rezultati desezoniranja pogoršani (postali *Severe* ili *Uncertain*), poželjno je korigovati pojedinačne specifikacije i sve izmene zapamtiti u postojećem *Processing*-u.



Slika 5.53: Potvrda Refresh-a

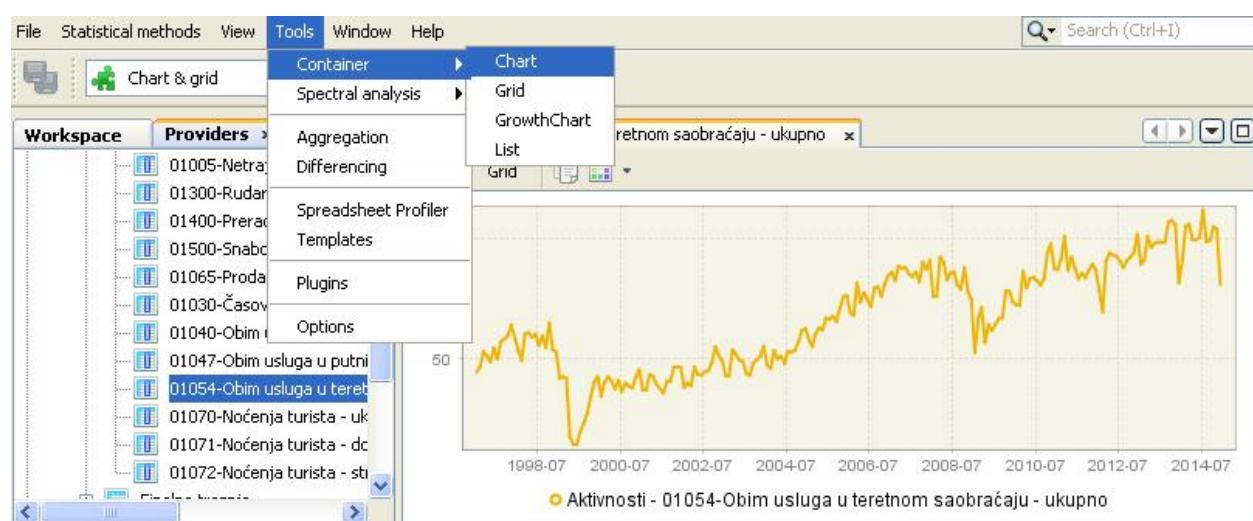
## 6 PRIMERI KORIŠĆENJA ALATA JDemetra+ (2.0.0)

### 6.1 Poređenje uticaja kreirane i *default* specifikacije

Prvi primer predstaviće način na koji bi korisnik, kada dobije sirovu seriju, najpre trebalo da joj pristupi, odakle da počne i kako da na osnovu grafičkih interpretacija i logičkih smernica prepostavi korake u odabiru specifikacije. Radi poređenja, posmatraće se i automatska specifikacija, koju za datu seriju bira softver. Analizom rezultata biće predstavljeni neki od načina koji mogu pomoći u odabiru boljeg modela, a samim tim i objašnjenja eventualno loših procena. Naravno, korisnik ima slobodu da u donošenju odluka iskoristi poznavanje same prirode serije i iskustvo u ranijim postupcima desezoniranja, jer to u svakom slučaju doprinosi kvalitetu obrade.

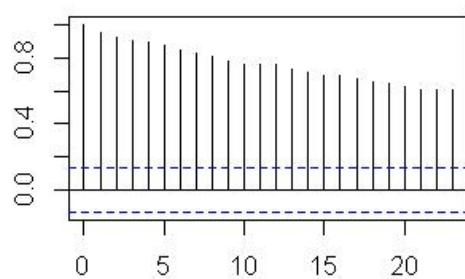
Vremenska serija koja će biti korišćena u ove svrhe predstavlja *obim usluga u teretnom saobraćaju* i duga je 18 godina (januar 1997 - decembar 2014).

Prvi korak, nakon importovanja željene serije, je najjednostavnija grafička predstava sirovih (originalnih, neobrađenih) podataka (*Tools → Container → Chart*):



Slika 6.1: Grafički prikaz serije

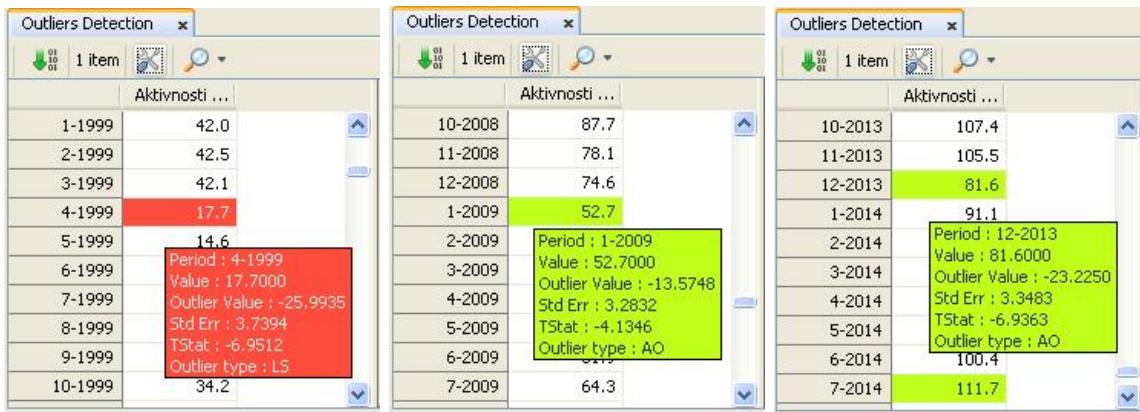
Prva impresija je da se posmatra serija uzlaznog trenda, nestacionarna u sredini, a, ako se izuzmu autolajeri, stacionarna u varijansi. Dakle, grafik indicira da serija ne mora biti logaritamski transformisana, tj. da je aditivna, ali da bi se postigla stacionarnost mora se izvršiti prva differencija. Pozitivna i lagano opadajuća autokorelaciona funkcija originalne serije potvrđuje nestacionarnost (slika 6.2).



Slika 6.2: ACF originalne serije

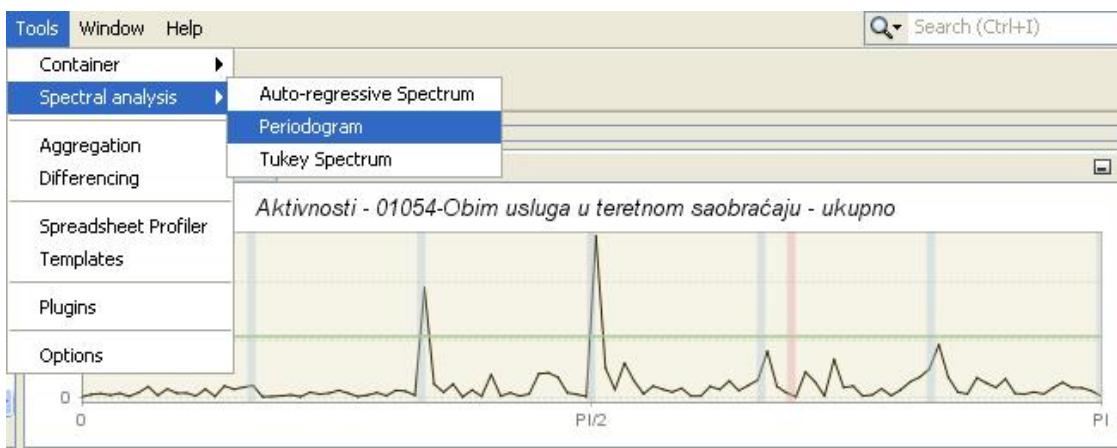
Sa grafika originalne serije takođe se mogu uočiti potencijalni autolajeri. Iskakanje sa putanjem serije uočava se u aprilu 1999. godine, januaru 2009., februaru 2012., decembru 2013. i julu i decembru 2014. godine. Tačan mesec i vrednost promenljive u toj vremenskoj tački

dobijaju se prostim prevlačenjem miša po grafiku. Pretpostavka je da je prvi autlajer *level shift*, zbog primetnog pada i izazvane dugoročne promene toka serije, dok su ostali, sudeći po grafičkoj predstavi, verovatno aditivni, jer se pojavljuju u izolovanim tačkama i regularan tok serije se brzo ponovo uspostavlja. Diskutabilna je jedino promena u januaru 2009. godine. No, da bi se utvrstile pretpostavke (ako već ne postoje ekonomske potvrde ovakvih iskakanja), JDemetra+ nudi mogućnost detektovanja autlajera još u prvoj fazi analize, pre procesa desezoniranja (*Statistical methods → Anomaly Detection → Outliers Detection*). Na slici 6.3 prikazani su neki od njih, ali su ovom metodom sve pretpostavke potvrđene.



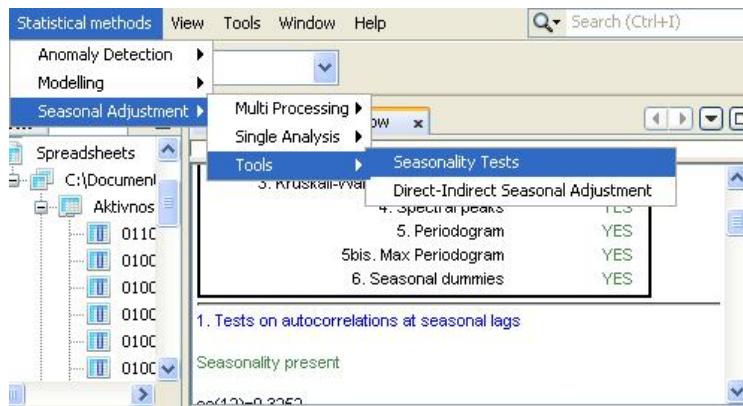
Slika 6.3: Detektovanje autlajera

Još jedna bitna informacija tiče se prisustva sezonalnosti. Korisnik mora da utvrdi da ima smisla vršiti desezoniranje, a s tim u vezi je posmatranje periodograma (*Tools → Spectral analysis → Periodogram*) (slika 6.4). Vrhovi na plavim linijama, iznad zelene, ukazuju na prisustvo sezonalnosti, a to što na crvenoj liniji toga nema znak je da na seriju ne utiče efekat kalendara. Ako se osvrnemo na prirodu serije, jasno je da za promet u saobraćaju zaista nikakvu bitnu ulogu ne igraju nacionalni praznici, tj. usluge transporta ni u kom slučaju tokom godine ne staju.



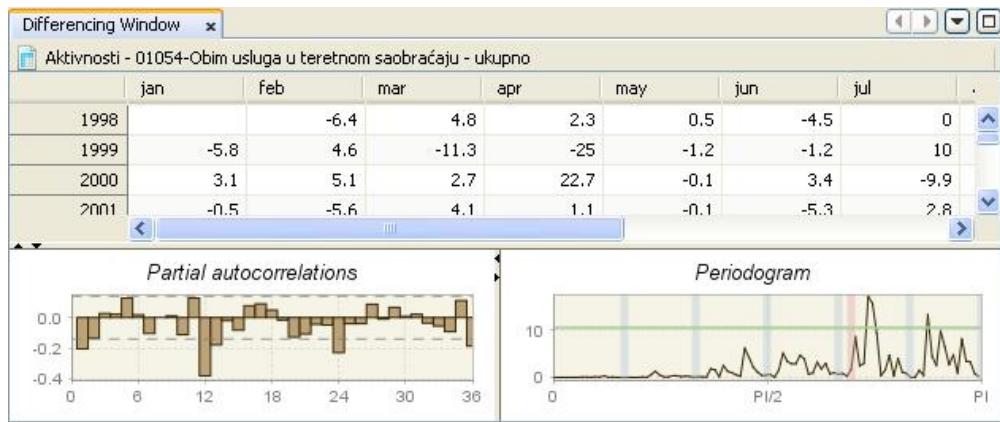
Slika 6.4: Periodogram originalne serije

Iako periodogram jasno pokazuje da je posmatrana serija sezonska, ako korisnik želi u predfazi procesiranja statističkim testovima to da potvrди, i to je omogućeno (*Statistical methods → Seasonal Adjustment → Tools → Seasonality Tests*) (slika 6.5). Prvo je data sumarna procena, a onda za svaki test pojedinačno objašnjenje. U ovom slučaju, svi su dali potvrdu da sezonalnost jeste prisutna.



Slika 6.5: Testovi za proveru prisustva sezonalnosti

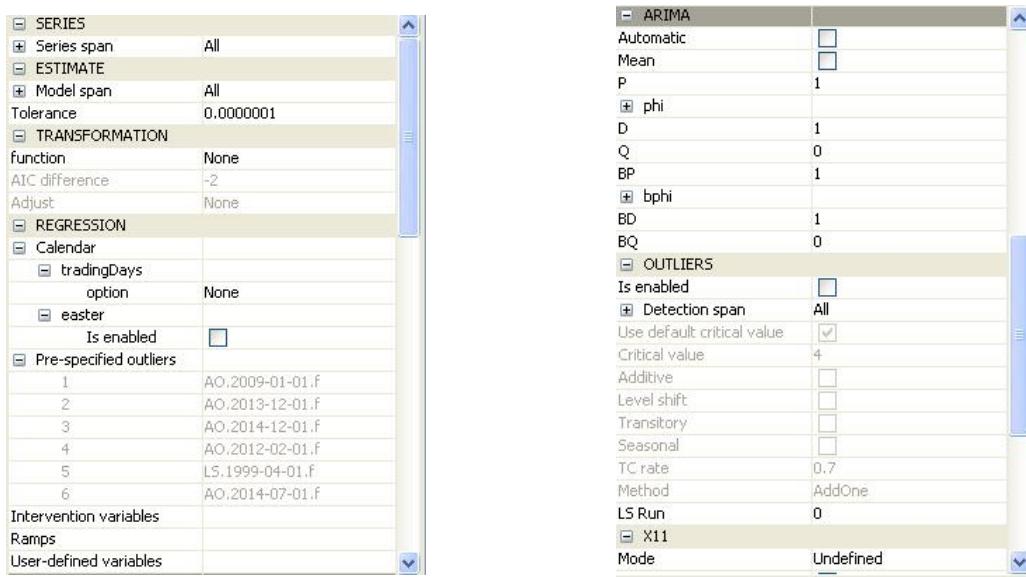
S obzirom na to da je serija nestacionarna i da se postizanje stacionarnosti vrši diferenciranjem, treba promotriti i tu fazu analize (*Tools → Differencing*) (slika 6.6). U novom prozoru predstavljene su razlike (naravno, prvi broj je za februar 1998. godinu), parcijalna autokorelaciona funkcija prvih razlika i periodogram. Jasno je da je posle diferenciranja izgubljeno prisustvo sezonalnosti, a neposredno sa PACF grafika mogu se naslutiti eventualni parametri ARIMA modela. Laički govoreći, kako je jedan pravougaonik van granica i on se ponavlja na dvanaest meseci, ARIMA model može biti tipa  $(1, 1, 0)$ , a zbog utvrđene sezonalnosti originalne serije i ponavljanja staze PACF grafika, za sezonski deo takođe se može pretpostaviti da je  $(1, 1, 0)$ . Naravno, ovo su sve procene koje nikako nisu potvrđene, te se uzimaju sa izvesnom rezervisanošću.



Slika 6.6: Prva differenca originalne serije

Kada korisnik završi sa sirovom analizom serije, sledeći korak je pravljenje sopstvene specifikacije. Nakon pokretanja procesa desezoniranja, ako rezultati nisu zadovoljavajući, specifikaciju treba korigovati dok kvalitet obrade ne bude zadovoljavajući. Prepostavke koje bi u ovom konkretnom primeru eventualno bile od koristi, sudeći po prethodnom posmatranju originalnih podataka, su:

- transformacija podataka - *None* (jer je serija aditivna);
- kalendar - bez efekta kalendar (trading/working dana, praznika, prestupne godine);
- autlajeri - unapred definisati uočene potencijalne autlajere;
- ARIMA model - samostalno definisati prepostavljeni model  $((1, 1, 0)(1, 1, 0))$ ;

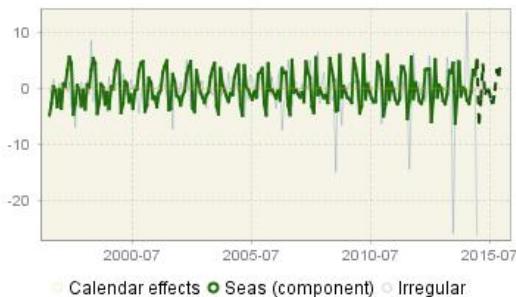


Slika 6.7: Kreirana specifikacija

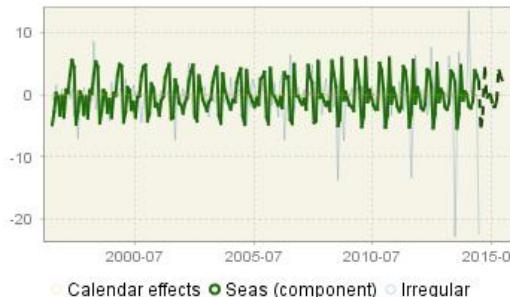
Sa ovako kreiranom specifikacijom pokreće se proces desezoniranja (i to *Single Processing*, jer se radi samo o jednoj seriji). Takođe, posmatra se i drugi proces, sa potpuno automatskom specifikacijom (bez ikakvih izmena). Da bi korisnik mogao da odluci koji od ponuđenih modela obrade je bolji, prvo treba da poredi sumarne rezultate, a zatim da po nekim kriterijumima presudi. Na primer, mi ćemo koristiti Akaikeov kriterijum i posmatrati grafike sezonskih komponenti (onoga što se iz serije odstranjuje). Proces koji ima manju vrednost Akaikeovog kriterijuma smatra se boljim, a takođe, onaj koji ima stabilniju (po varijansi i očekivanju) sezonsku komponentu pokazuje tendenciju ka boljem odabiru. Zatim ćemo pokušati da nađemo razlog i objašnjenje za ono što je zapravo dovelo do boljstva u kvalitetnijem slučaju.

Sumarni rezultati automatske obrade pokazuju da softver nije zabeležio efekte kalendara (ni *trading days effects*, ni *easter effects*), kao ni potrebu za logaritamskom transformacijom, što potvrđuje za sad odluke donešene u slučaju kreiranja specifikacije. Takođe, program je, kako smo i mi grafički uočili, pronašao šest statistički značajnih autolajera. Ukupan kvalitet obrade u oba slučaja ocenjen je kao *Good*, s tim što su neki niži nivoi procene kvaliteta označeni crvenom bojom u slučaju korišćenja sopstvene specifikacije. To, u prvi mah, indicira da je automatska obrada bolja.

Grafici sezonskih komponenti su vrlo slični i teško se može doneti zaključak o sabilnjoj komponenti (slika 6.8), mada se može primetiti da manje varira oko nule sezonska komponenta u slučaju automatske specifikacije. Uočava se još da se staze kretanja za nijansu razlikuju i da je to naročito izraženo na kraju serije i u prognozi.



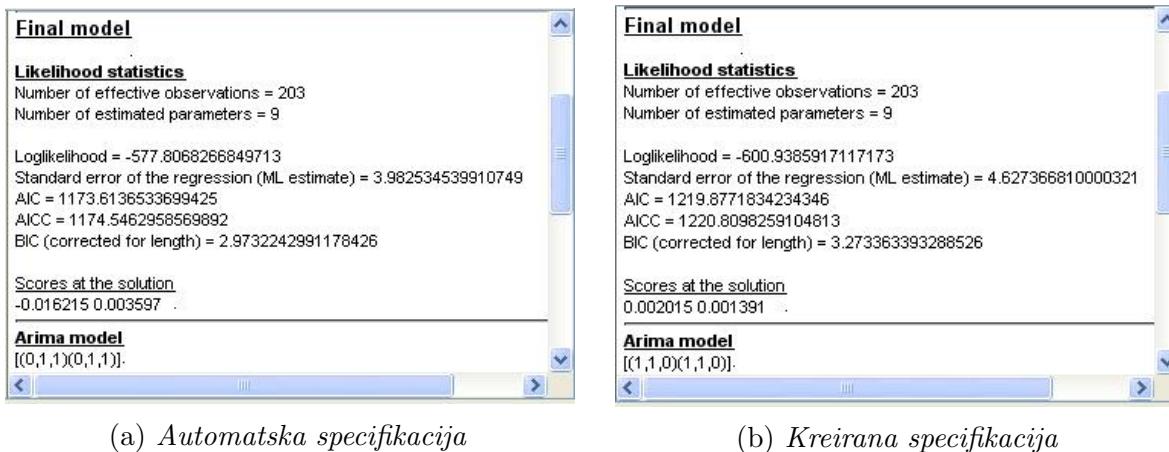
(a) Automatska specifikacija



(b) Kreirana specifikacija

Slika 6.8: Poređenje grafika sezonskih komponenti

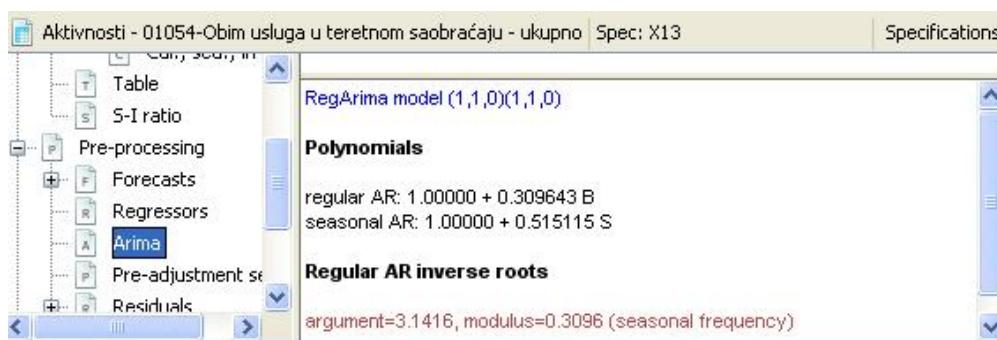
Stoga, da bi se potvrdila pretpostavka da je automatska specifikacija kvalitetnija, posmatra se Akaikeov kriterijum:



Slika 6.9: Poređenje vrednosti Akaikeovog kriterijuma

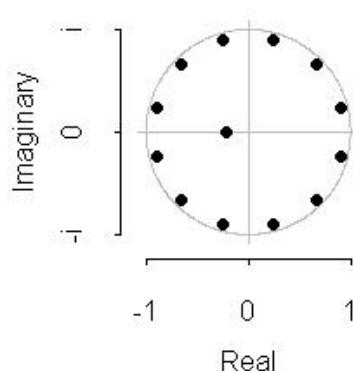
Zaista, u sučaju automatske obrade, manje vrednosti Akaikeovog kriterijuma i standardne greške regresije potvrda su da ovakva specifikacija bolje pristaje posmatranoj vremenskoj seriji.

S obzirom na to da se odluka o transformaciji, kalendarskim efektima i autolajerima poklopila sa softverom, uzrok razlika u kvalitetu desezoniranja treba tražiti u odabiru modela. Procena na osnovu grafika bila je da bi datoj vremenskoj seriji mogao da pristaje model  $(1, 1, 0)(1, 1, 0)$ , međutim ako se u rezultatima (*Pre-processing* → *Arima*) proveri Arima model, uočava se crveno obaveštenje, koje signalizira nekorektnost sa stohastičkom sezonalnošću (slika 6.10).



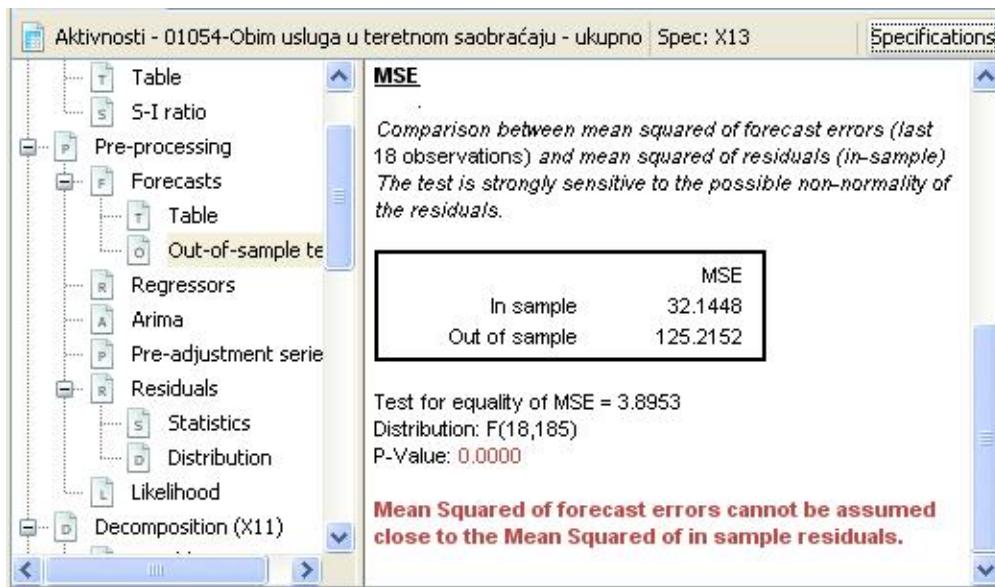
Slika 6.10: Arima model za kreiranu specifikaciju

Naime, postoji mogućnost da nakon desezoniranja, serija i dalje nije stacionarna i da ima jedinične korene na sezonskim frekvencijama. Kreirala sam program u statističkom softveru *R*, koji iscrtava sliku 6.11 i ovakav zaključak potvrđuje. Najbolje se prisustvo sezonalnosti ogleda u autokorelaciji reziduala. Istina, rezultati *Ljung-Box* i *Box-Pierce* testa za datu obradu alarmiraju crvene rezultate za testiranje nezavisnosti reziduala. To su one crvene naznake u nižim nivoima rezultata, spomenute na početku, koje nisu uvek bitne za globalnu ocenu desezoniranja, ali jesu ako se bira bolji od dva modela.



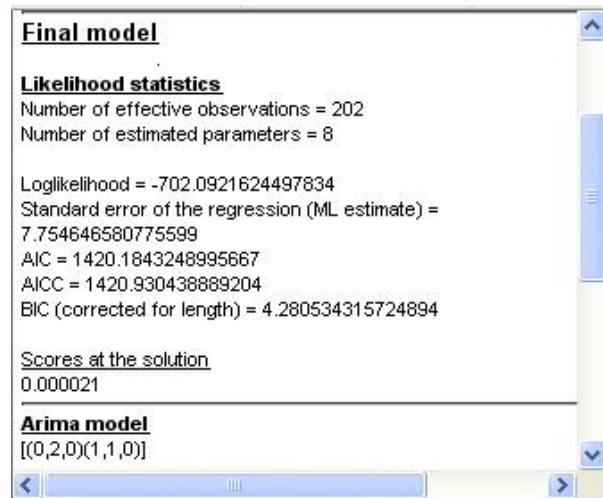
Slika 6.11: Inverzni AR korenji

Još jedan simptom koji ukazuje na jedinične korene, tj. na nestacionarnost u seriji, je prognoza modela, koja se obično besmisleno ponaša. Zaista, u slučaju atomatske specifikacije, interval poverenja za prognozu je uži, što ukazuje na veću preciznost, a test koji poredi srednje kvadratnu grešku prognoze sa rezidualima indicira da nisu saglasni, jer je model uticao i na (ne)normalnost reziduala.

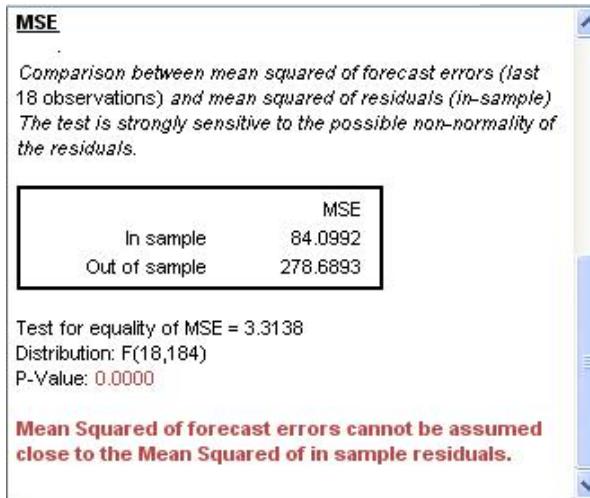


Slika 6.12: *Out-of-sample test za model  $(1, 1, 0)(1, 1, 0)$*

Kad postoje jedinični koreni u AR delu modela (statistički značajno blizu jediničnog kruga ili na njemu), trebalo bi redukovati parametar  $p$ , a povećati red diferenciranja za jedan. Ako se u kreiranoj specifikaciji promeni samo Arima model na  $(0, 2, 0)(1, 1, 0)$  zaista se gubi upozorenje o prisustvu jediničnih korena. Međutim, Akaikeov kriterijum je sada daleko lošiji (slika 6.13), a rezultat testa za proveru saglasnosti grešaka prognoze sa rezidualima takođe je crveno upozorenje. Stoga, ostaje zaključak da je u ovom primeru bolji kvalitet rezultata dala specifikacija preporučena softverom.



Slika 6.13: *Akaikeov kriterijum za model  $(0, 2, 0)(1, 1, 0)$*



Slika 6.14: *Out-of-sample test za model  $(0, 2, 0)(1, 1, 0)$*

## 6.2 Uticaj autlajera na liniju trenda

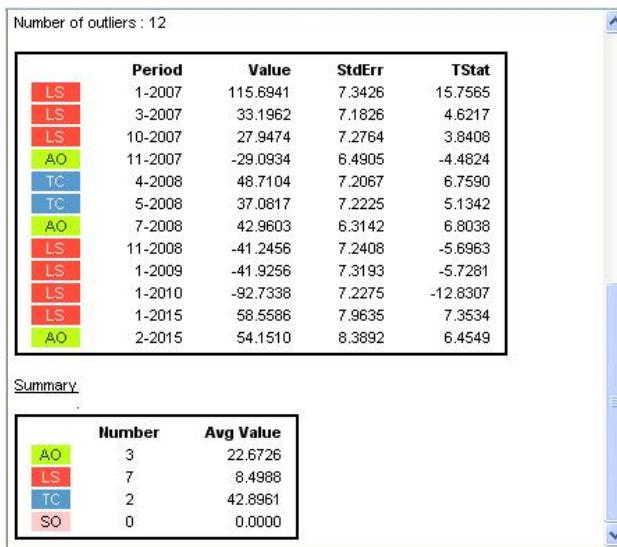
Autlajeri nose neku informaciju. Bilo da je njihova prisutnost statistički manje ili više značajna, oni govore o određenom tipu promena u toku serije, te se njihov uticaj ne sme zanemariti. JDemetra+, kako je već spomenuto, prepoznaće tri tipa ekstremnih vrednosti - AO, TC i LS, od kojih LS autlajeri direktno utiču na liniju trenda, a AO i TC autlajeri na iregularnu komponentu. Međutim, kako je  $sa = i \cdot t$ , povećanje/smanjenje iregularne ili trend komponente indirektno utiče i na onu drugu.

Posmatraćemo seriju iz sektora Spoljna trgovina, *Uvoz robe - neklasifikovano po BEC<sup>13</sup> nameni*), koja je dužine 11 godina (januar 2004 - februar 2015).



Slika 6.15: Grafik posmatrane serije

Ova serija je vrlo iregularna i neposredno sa grafika se uočava da ima veliki broj autlajera. Korišćenjem opcije *Outlier detection* dobija se izlaz sa slike 6.16. Dakle, serija ima tri aditivna, sedam LS i dva TC autlajera, što znači da će odluka o isključivanju nekih od njih, rezultate procesa desezoniranja znatno izmeniti.

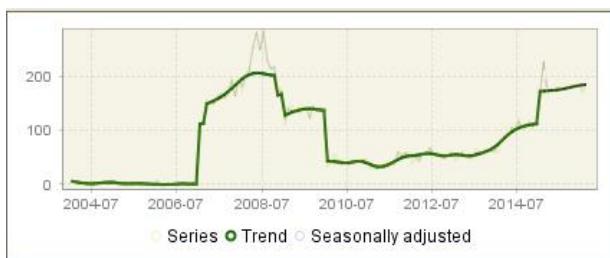


Slika 6.16: Detektovani autlajeri posmatrane serije

Ako se posmatra linija trenda desezonirane serije u slučaju kada su uključene sve tri vrste autlajera (slika 6.17), a poznajući prirodu njihovog uticaja, jasno je da su oštiri vrhovi posledica dejstva LS autlajera, jer oni direktno utiču na liniju trenda. Stoga, isključivanjem LS autlajera iz analize, očekuje se glatka linija trenda, što je i na slici 6.18 potvrđeno. Kako

13. BEC - klasifikacija (engl. *Classification by Broad Economic Categories*)

je i ideja da linija trenda pokazuje dugoročnu tendenciju kretanja serije, ukoliko ima puno LS autlajera, treba ih u većini slučajeva odstraniti, naročito ako je serija, kao sada, iregularna i njihov uticaj ovako izražen.



**Slika 6.17:** Trend linija kada je uključen uticaj svih autlajera

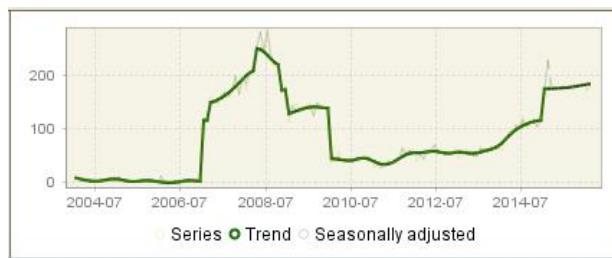


**Slika 6.18:** Trend linija bez uticaja LS autlajera

S druge strane, odstranjivanje samo AO ili samo TC autlajera je samo gubitak informacije, a linija trenda bi ostala sa oštrim vrhovima (slike 6.19 i 6.20), pa se to ne preporučuje. Program je detektovao AO autlajere u novembru 2007. godine i februaru 2015. godine, pa se na slici 6.19 tačno vidi veći prelom na tim tačkama, a registrovani TC autlajeri su u aprilu i maju 2008. godine, pa je na slici 6.20 u tim tačkama jedina razlika u odnosu na sliku 6.17.



**Slika 6.19:** Trend linija bez uticaja AO autlajera



**Slika 6.20:** Trend linija bez uticaja TC autlajera

Ako se slika 6.21 uporedi sa slikom 6.18, vidi se da bi se trend u ovom slučaju tačno razlikovao za tačke sa istaknutim AO i TC autlajerima, što bi naravno izvrnulo sliku i o iregularnoj komponenti, kao i o desezoniranoj seriji.



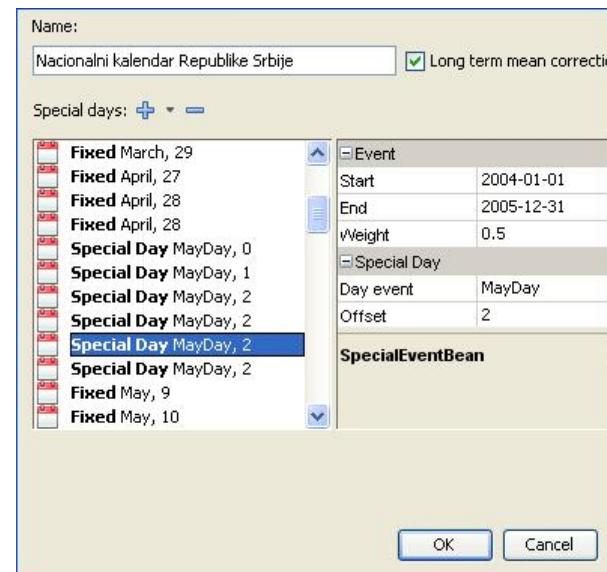
**Slika 6.21:** Trend linija bez uticaja svih autlajera

Da se nešto ne označi kao ekstremna vrednost (iako to jeste), već se računa kao deo normalnog toka serije, izvrnulo bi predstavu o dugoročnoj stazi vremenske serije, što bi pružilo neprecizne informacije. Stoga je uvek prva preporuka proveriti liniju trenda i ako autlajeri ne prave šljgate vrhove i nagle prelome, sve ih treba ostaviti.

### 6.3 Poređenje uticaja nacionalnog i *default* kalendara

Kao što je ranije spomenuto, uticaj efekata kalendara može biti izuzetno bitan za praćenje toka vremenske serije. Kalendar Republike Srbije vrlo je kompleksan, s obzirom na to da obuhvata nacionalne praznike koji datiraju od 1994. godine, s tim što su neki od njih 2007. ukinuti (npr. 28. mart - Dan državnosti ili 27. april - Dan Savezne Republike Jugoslavije, 9. maj - Dan pobede ili 7. jul - Dan ustanka naroda Srbije), a neki novi uvedeni (npr. 7. januar - Pravoslavni Božić ili 15. februar - Dan državnosti). Takođe, od 2012. godine uveden je kao državni praznik i 11. novembar - Dan primirja u Prvom svetskom ratu, a Dan državnosti praznuje se dva radna dana.

Umanjenim težinskim koeficijentima računaju se dani kada se praznik prenese na prvi radni dan, jer je kalendarski pao na vikend. Takođe, određeni ponderi uzimaju se i za najveće slave kod nas, jer je, na primer, za Svetog Nikolu znatno manji broj zaposlenih prisutan na svojim radnim mestima. Poseban režim se u JDemetri+ koristi za ocenjivanje uticaja pravoslavnog Uskrsa, jer se program služi Gregorijanskim kalendарom. Svake godine se poredi razlika u danima između katoličkog i pravoslavnog Uskrsa i taj se odnos beleži opcijom *Easter Related*. Zbog kompleksnosti, za potrebe ovog primera biće importovan već kreirani kalendar, po principu objašnjrenom u poglavlju 5.3, a prikazan na slici 6.22.



Slika 6.22: Nacionalni kalendar

U ovom primeru biće posmatrana serija iz sektora Platni bilans, *Rashodi za uvoz saobraćajnih usluga*, dužine 7 godina (januar 2007 - april 2014):



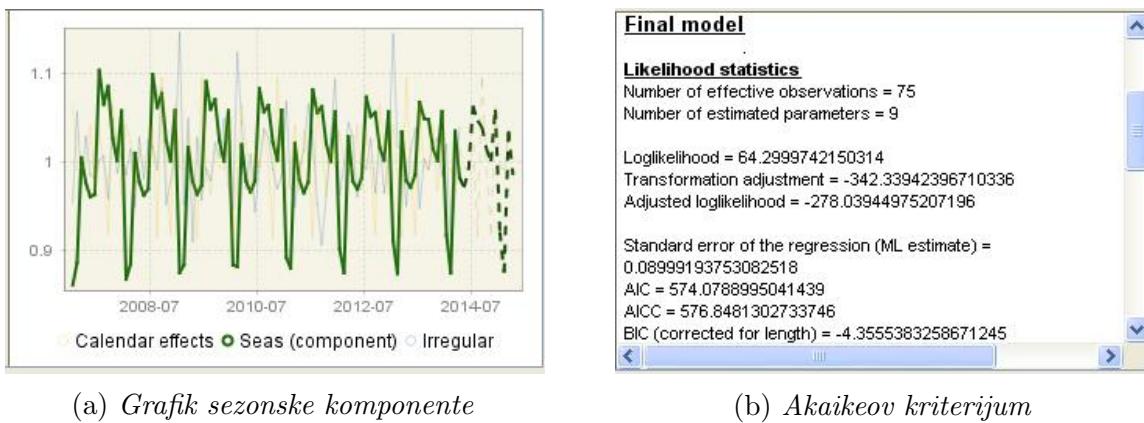
Slika 6.23: Grafik posmatrane serije

Sa periodograma (slika 6.24) se učava izraženi uticaj efekta kalendara, pa ima smisla uočavati razliku u kvalitetu modela kad se specifikacije razlikuju samo po kalendaru koji se koristi - nacionalni ili *default*.

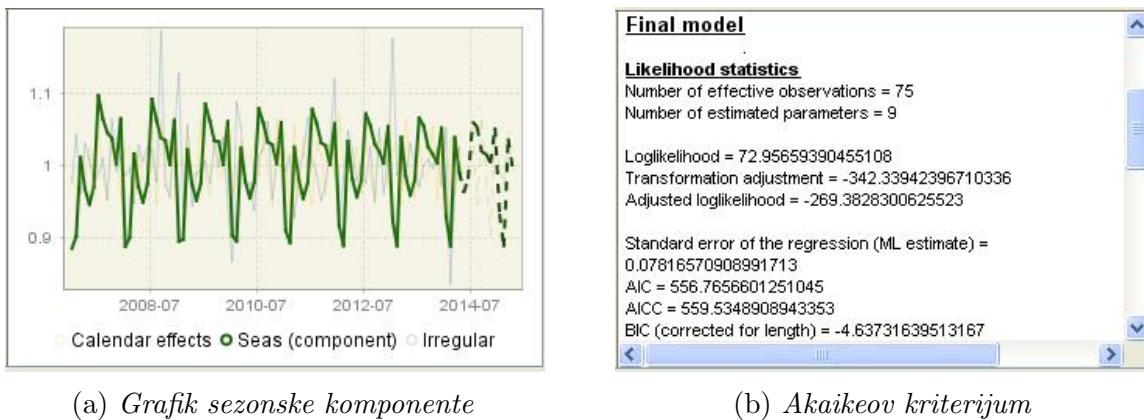


Slika 6.24: Periodogram posmatrane serije

Odluku o boljem modelu donećemo posmatranjem stabilnosti odstranjene sezonske komponente i Akaikeovog kriterijuma u oba slučaja.



Slika 6.25: Rezultati uticaja default kalendarja



Slika 6.26: Rezultati uticaja nacionalnog kalendarja

Iz priloženog se može primetiti da u slučaju korišćenja nacionalnog kalendarja sezonska komponenta manje varira oko jedinice, tj. da je stabilnija, kao i da su vrednosti Akaikeovog i Bajesovog informacionog kriterijuma i standardne greške regresije manji. Ovo je dovoljno da se zaključi da se reprezentativniji model dobija kada se za efekat kalendarja računa nacionalni kalendar, što se i moglo očekivati.

## 6.4 Poređenje X13 i TramoSeats metoda kod nesezonske serije

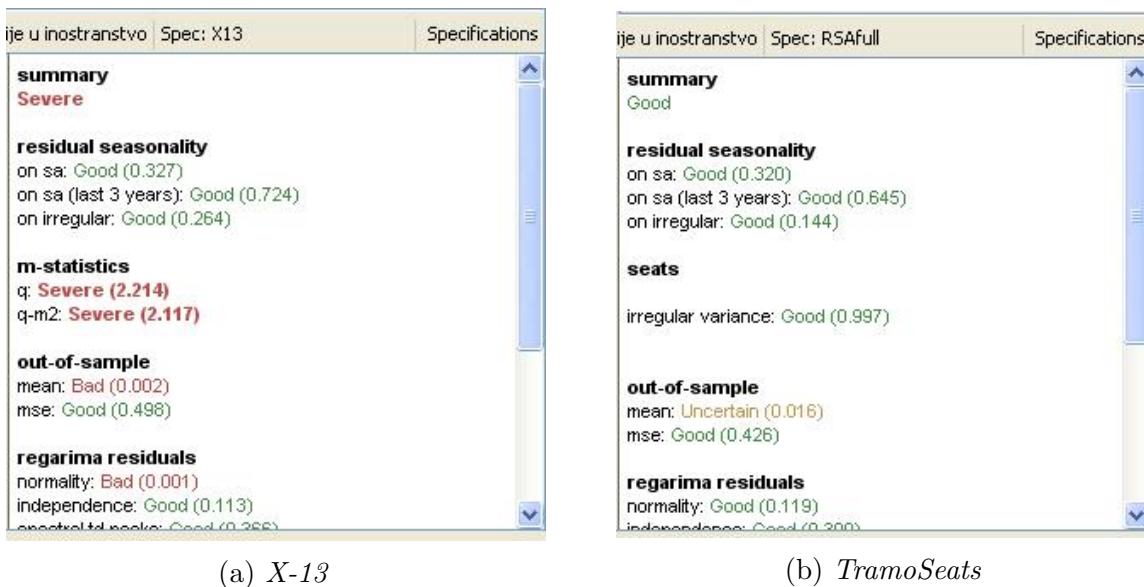
Osnovna razlika između X-13 i TramoSeats metoda je što X-13 uvek ocenjuje sezonsku komponentu, jer je zasnovana na filterima (engl. *filter based*), dok TramoSeats, ukoliko program proceni nepostojanje sezonalnosti, seriju tako i tretira, jer je orijentisan ka modelu (engl. *model based*). Stoga i važi nezvanično, ali potvrđeno pravilo, da je nesezonske serije bolje obradivati TramoSeats metodom.

Vremenska serija iz sektora Platni bilans, *Direktne investicije u inostranstvo*, duga sedam godina (januar 2007 - april 2014), poslužice za interpretaciju razlike koja nastaje kada se koriste dve različite metode.



Slika 6.27: Grafik posmatrane serije

Jasno je da je serija vrlo iregularna, da na početku ima veliki skok, koji potpuno remeti tok serije, a takođe, neposredno sa grafika se može zaključiti da serija nema ustaljenu formu kretanja, te da je efekat sezonalnosti odsutan. S obzirom na to da akcenat neće biti na korišćenoj specifikaciji, u oba slučaja biće korišćena automatska specifikacija, i to RSA5c za X-13, a RSAfull za TramoSeats.

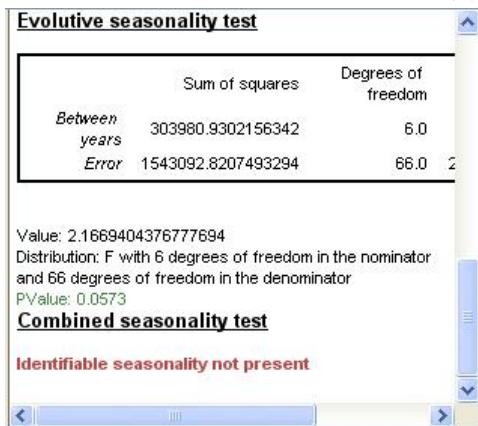


Slika 6.28: Poređenje sumarnih rezultata obrade

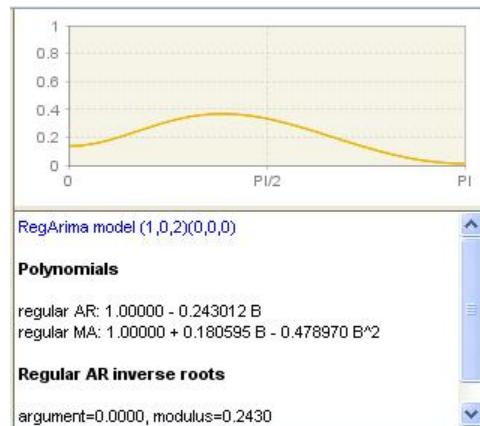
Već na prvi pogled se primećuje da su sumarni rezultati obrade X-13 metodom daleko lošiji. Potvrde odsutnosti sezonalnosti su periodogram, testovi za proveru prisustva sezonalnosti i tip automatski odabranog ARIMA modela. Sa slike 6.29 vidi se da periodogram nema špicave na plavim linijama, da kombinovani test (a i ostali testovi) ne detektuje sezonalnost, kao i to da su parametri sezonskog dela ARIMA modela svi nula ((1, 0, 2)(0, 0, 0)).



(a) Periodogram



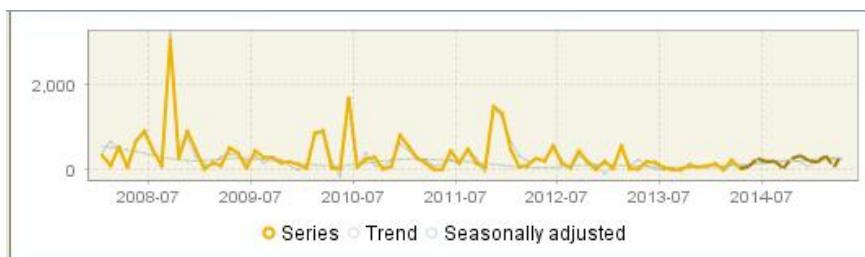
(b) Kombinovani test za sezonalnost



(c) ARIMA model

Slika 6.29: Potvrde odsustva sezonalnosti

Jedan od načina da se dobije bolji opšti rezultat obrade X-13 metodom je da se pokuša sa skraćivanjem serije (engl. *Series span*). Ako se odstrani prvi deo date serije (posmatra se serija od januara 2008, dobija se:



Slika 6.30: Grafik skraćene serije

U startu se odričemo dela informacije, što je, naravno, za praćenje toka serije vrlo pogubno, naročito jer ova serija sama po sebi nije velike dužine. S druge strane, da bi se kvalitet rezultata poboljšao, makar do *Uncertain* (jer sa ovako iregularnom serijom to je najbolje što se može očekivati), treba povećati sezonski filter, jer to izglačava seriju kad je sezonalnost iregularna. Softver pokazuje da je koristio  $S3 \times 9$ , pa ručno postavljamo  $S3 \times 15$ . Ukupan kvalitet će se zaista (po izlazu koji softver nudi) popraviti (slika 6.31), međutim, dugačak filter treba primenjivati samo ukoliko je serija dovoljno duga, što ovde nije slučaj, pa je samo poboljšanje diskutabilno. Takođe, što je duži filter, to se prognozi treba manje verovati, pa uvid u tok serije postaje, globalno, vrlo loš. Iz svih ovih razloga, u ovakvim slučajevima najbolje je koristiti *TramoSeats*.



Slika 6.31: Opšti rezultat

Ako se sada posmatra rezultat *TramoSeats* obrade, na slici 6.32 uočava se nepostojanje grafika sezonske komponente, iz prostog razloga što se ovde i ne računa, jer testovi pokazuju da sezonalnost nije prisutna. Nema ni prisustva efekta kalendara, pa se ceo tok serije pripisuje uticaju iregularne komponente.



Slika 6.32: Main results → Cal., sea., irr (*TramoSeats*)

S druge strane, u tablici sa slike 6.33 jasno se uočava da su originalna i desezonirana serija identičnih vrednosti. Kako je model multiplikativni (jer je program izabrao kao najpodesniju logaritamsku transformaciju), a sezonska komponenta ocenjena kao nepostojeća (svuda 1), to je desezonriana serija, koja se dobija po obrascu  $sa = y/s$ , poistovećena sa originalnom. Ovo je jasan dokaz da *TramoSeats* metoda za desezoniranje uvek opominje korisnika da nema potrebe desezonirati seriju ako već sezonski efekat ne postoji.

	Series	Seasonall...	Trend	Seasonal	Irregular
1-2007	61.1	61.1	901.912	1	0.068
2-2007	32.4	32.4	889.727	1	0.036
3-2007	318.8	318.8	881.297	1	0.362
4-2007	294.3	294.3	874.275	1	0.337
5-2007	6.7	6.7	869.556	1	0.008
6-2007	22,536.6	22,536.6	866.821	1	25.999
7-2007	295	295	861.426	1	0.342
8-2007	662.7	662.7	853.264	1	0.777

Slika 6.33: Main results → Table (*TramoSeats*)

## 7 ZAKLJUČAK

Desezoniranje, kao i, generalno, analiza vremenskih serija, sa sobom nose rizike jednako koliko i eventualnu mogućnost boljeg razumevanja prirode podataka. Modeliranje samo po sebi vrlo je kompleksan proces i samim tim što se najviše oslanja na intuiciju i iskustvo analitičara, često je prezasićeno subjektivnošću. Međutim, korisno je ovde citirati *George E.P. Box*-a, koji navodi da su svi modeli pogrešni, ali da su neki sigurno korisni. Naravno, sve i ako se treba oslanjati na aproksimacije, ako je problem dobro definisan, rezultat je kvalitetniji nego da se apsolutno tačan odgovor nalazi za pogrešno pitanje.

Metode i tehnike za manipulaciju podacima, u cilju boljeg razumevanja ekonomskih i socijalnih indikatora, neprestano su u razvojnom procesu. Time nivo statističke istine stalno raste. Korisnici su uvek upozorenici da sa određenom dozom rezervisanosti treba upotrebljavati prognoze i predikcije, jer su proizvod prošlosti i sadašnjosti, koje su izložene konstantnim promenama, na koje se ne utiče. *Ronald Coase* navodi da ako se podaci, uslovno rečeno, dovoljno muče, priroda će uvek priznati. Istina je da analitičari i statističari imaju sve slobode da luksuz pojavljivanja 'čudnih' događaja tretiraju kako žele, ali *Mark Twain*, u prilog ovome, kaže: *Sve generalizacije su netačne, uključujući i ovu*, što znači da se u moru procena uvek nađe dovoljno dobra.

Najlepše u procesu modeliranja je kada vas slika natera da primetite ono što nikad ne biste očekivali. Upravo su visoke ocene kvaliteta poslovanja indikator dobrog radnog procesa, bilo da na vreme otkrivaju potencijalne rizike ili su pokazatelj napretka. Vreme je, kao i procene, relativno, ali je neophodno za razumevanje realnih tokova događaja. Stoga, bez obzira na uzdržanost u smislu apsolutne tačnosti rezultata, analiza vremenskih serija je široko primenjena i neizostavna oblast u svim granama privredne i društvene delatnosti. U prilog tome, *Stephen Senn* konstatiše da 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.

Naposletku, kada dodu neki jesenji, sumorni dani i čini se da je teško naći smisao - treba se setiti da je, desezonirano gledano, sreća sveprisutna i statistički najznačajnija, kao dugoročna tendencija iskoruščavanja lepote koju nosimo u sebi.

*Ja nisam autlajer, samo još nisam našao svoju raspodelu. (Ronan Conroy)*

## 8 LITERATURA

- [1] Ross Ihaka, *Time Series Analysis*, Statistics Department, University of Auckland, 2005.
- [2] George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel, *Time Series Analysis - Forecasting and Control (Fourth Edition)*, Statistics Department, United States of America, 2008.
- [3] Ruey S. Tsay, *Analysis of Financial Time Series (Second Edition)*, University of Chicago, Graduate School of Business, 2005.
- [4] Chair of Statistics, University of Würzburg, *A First Course on Time Series Analysis*, March 20, 2011.
- [5] William W.S. Wei, *Time Series Analysis - Univariate and Multivariate Methods (Second Edition)*, Department of Statistics, The Fox School of Business and Management, Temple University, 2006.
- [6] Gebhard Kirchgässner, Jürgen Wolters, *Introduction to Modern Time Series Analysis*, Springer-Verlag, Berlin, Heidelberg, 2007.
- [7] Zlatko J. Kovačić, *Analiza vremenskih serija*, Ekonomski fakultet, Univerzitet u Beogradu, 1995.
- [8] Daniel Pena, George C. Tiao, Ruey S. Tsay, *A Course in Time Series Analysis*, Statistics Department, Universidad Carlos III de Madrid and University of Chicago, 2001.
- [9] Chris Chatfield, *The Analysis of Time Series - an Introduction (Fifth Edition)*, Reader in Statistics, The University of Bath, United Kingdom
- [10] Zorica Mladenović, Aleksandra Nojković, *Primenjena analiza vremenskih serija*, Centar za izdavačku delatnost ekonomskog fakulteta u Beogradu, 2012.
- [11] Danijela Rajter-Ćirić, *Verovatnoća*, Univerzitet u Novom Sadu, Prirodno-matematički fakultet u Novom Sadu, Novi Sad, 2009.
- [12] Danijela Rajter-Ćirić, *Beleške sa kursa Stohastička Analiza*, Novi Sad, 2012.
- [13] Peter Thomson, Tohru Ozaki, *Transformation and Seasonal Adjustment*, Statistics Research Associates Ltd (Wellington, New Zealand), Institute of Statistical Mathematics (Tokyo, Japan), 1998.
- [14] Dominique Ladiray, Benoit Quenneville, *Seasonal Adjustment with the X-11 Method*, Eurostat and Statistics Canada, 2001.
- [15] Methodologies and Working papers, *ESS Guidelines on Seasonal Adjustment*, Eurostat, Luxembourg, 2009.

- [16] Information paper, *An Introductory Course on Time Series Analysis - Electronic Delivery*, Australian Bureau of Statistics, 2005.
- [17] J.D. Hamilton, *Time Series Analysis*, Princeton University Press, New Jersey, 1994.
- [18] Hung T. Nguyen, Berlin Wu, *Fundamentals of Statistics with Fuzzy Data*, 2006.
- [19] Michele Manna, Romana Peronaci, *Seasonal Adjustment*, European Central Bank, 2003.
- [20] D. C. Montgomery, C. L. Jennings, M. Kulachi *Introduction to Time Series Analysis and Forecasting*, John Wiley and Sons, 2008.
- [21] Erika Foldesi, Peter Bauer, Beata Horvath, Beata Urr, *Seasonal Adjustment - Methods and Practices*, Hungarian Central Statistical Office, 2007.
- [22] Victor Gomez, Augustin Maravall, *Seasonal Adjustment and Signal Extraction in Economic Time Series*, Banco de Espana, 1997.
- [23] Peter Linde, *Seasonal Adjustment*, Statistics Denmark, 2005.
- [24] Bell W.R., Hillmer S.C., *Issues involved with the Seasonal Adjustment of Economic Time Series*, Journal of Business and Economic Statistics, 1984.
- [25] Sylwia Grudkowska, *JDemetra+ User Manual*, National Bank of Poland, October 2013.
- [26] United Nations, *Practical Guide to Seasonal Adjustment with Demetra+*
- [27] David F. Findley, Brian C. Monsell, William R. Bell, Mark C. Otto and Bor-Chung Chen, *New Capabilities and Methods of the X-12-ARIMA Seasonal Adjustment Program*, U.S. Bureau of the Census, 1998.
- [28] Hilmer, S.C and G.C. Tiao, *An ARIMA-Model-Based Approach to Seasonal Adjustment*, Journal of the American Statistical Association, 1982.
- [29] Estela Bee Dagum, *The X-II-ARIMA Seasonal Adjustment Method*, Statistics Canada, February 1980.
- [30] Jovan Mališić, Vesna Jevremović, *Slučajni procesi i vremenske serije*, Matematički fakultet, Univerzitet u Beogradu, 2008.