

СТРУЧНИ РАДОВИ
WORKING PAPERS

2013



Народна банка Србије
National Bank of Serbia

Радови у оквиру ове серије представљају резултат тренутних истраживања аутора и објављују се у циљу иницирања дискусије и добијања корисних сугестија за даљи рад аутора.

Working Papers describe research in progress by the author(s) and are published to encourage discussion and suggestions for future work.

СТРУЧНИ РАДОВИ
WORKING PAPER SERIES

26

**Детаљан приказ десезонирања серије
индустријске производње Србије методом
X-12-ARIMA**

Мирко Ђукић

Detailed description of X-12-ARIMA seasonal
adjustment of Serbian industrial production series

Mirko Đukić

© Народна банка Србије, март, 2013.
Доступно на www.nbs.rs

© National Bank of Serbia, March, 2013
Available at www.nbs.rs

За ставове изнете у радовима у оквиру ове серије одговоран је аутор и ставови не представљају нужно званичан став Народне банке Србије.

The views expressed in the papers constituting this series are those of the author, and do not necessarily represent the official view of the National Bank of Serbia.

Директорат за економска истраживања и статистику

НАРОДНА БАНКА СРБИЈЕ

Београд, Краља Петра 12,

Тел.: (+381 11) 3027 100

Београд, Немањина 17,

Тел.: (+381 11) 333 8000

www.nbs.rs

Directorate for Economic Research and Statistics

THE NATIONAL BANK OF SERBIA

Belgrade, 12 Kralja Petra Street,

Telephone: (+381 11) 3027 100

Belgrade, 17 Nemanjina Street,

Telephone: (+381 11) 333 8000

www.nbs.rs

Детаљан приказ десезонирања серије индустријске производње Србије методом X-12-ARIMA

Мирко Ђукић

Апстракт: У овом раду детаљно смо представили кораке десезонирања серије индустријске производње Србије за период 1994–2012. методом X-12-ARIMA. У фази пред-прилагођавања серију смо „прочистили“ од највећих аутоматски оцењених искока (*RegARIMA* моделом), као што су трајна промена нивоа за НАТО бомбардовање, и привремена промена за петооктобарске политичке промене итд, као и од „рампе“ за почетак светске економске кризе, коју смо сами укључили у регресију. Поред тога, *RegARIMA* моделом оценили смо да један радни дан мање у месецу, због празника или викенда, смањује производњу за 1,5%, а ти ефекти су такође (привремено) искључени из серије, пре уласка серије у X-11 фазу. Оцењена финална сезонска компонента серије у X-11 фази показује смањене сезонске осцилације у последњих десетак година, тј. све нижи сезонски фактор за јесењо-зимске и све виши за летње месеце. Серију индустријске производње смо поред десезонирања и „декалендарисали“, тј. искључили ефекте и броја радних дана и празника, чиме се добија реалнија слика о кретању у индустрији. Десезонирана серија садржи све искоке, с тим што се оцењени привремени искоци укључују у ирегуларну компоненту, а трајни искоци у тренд. У анализама ревизије јануар се испоставио као месец који је „најмање поуздан“, тј. као месец чија се десезонирана вредност највише ревидира додавањем нових података.

Кључне речи: X12-ARIMA, десезонирање, индустријска производња Србије

JEL Code: C22

Захваљујем Милану Недељковићу, Бранку Урошевићу, Бранку Хинићу и осталим колегама из Народне банке на корисним сугестијама, као и проф. Зорици Младеновић, код које сам одбранио магистарски рад на тему X-12-ARIMA.

Detailed description of X-12-ARIMA seasonal adjustment of Serbian industrial production series

Abstract: In this paper we seasonally adjust, step by step, monthly series of Serbian industrial production for the period 1994–2012, using the X-12-ARIMA method. In the pre-adjustment phase, we corrected the series for the automatically detected (using *RegARIMA*) effects of outliers, such as the level shift for NATO bombing, and the temporary change for the political changes of 5 October 2000 etc, as well as the (user-specified) ramp effect for the beginning of the global financial crisis. Using *RegARIMA* we also estimated that one working day less, because of weekends or holidays, reduces industrial production by 1.5%. These effects were also excluded from the series at this stage, before entering the next, X-11 phase. The final seasonal component in the X-11 phase, shows declining seasonal fluctuations of the series over the last 10 years, that is a lower seasonal factor for winter-autumn months and higher for summer months. Besides adjusting for the seasonal factor, we also adjusted the series for the calendar effect (number of working days and holidays), thus providing a more realistic picture on the movements in industry. Seasonally adjusted series contains all the outliers, with temporary ones being included in the irregular component, and permanent ones in the trend. The revisions analysis showed January as the least “reliable” month, that is the month for which seasonally adjusted data are revised mostly when new data are added to the series.

Key words: X-12-ARIMA, seasonal adjustment, industrial production

JEL Code: C22

Special thanks to Milan Nedeljković, Branko Urošević, Branko Hinić and other colleagues in the National bank of Serbia for useful suggestions, as well as prof. Zorica Mladenović, my master-thesis mentor on the topic of X-12-ARIMA.

Садржај:

1. Увод.....	1
2. Сврха и фазе процеса <i>X-12-ARIMA</i> метода	3
3. Предприлагођавање: <i>regARIMA</i> модел.....	4
3.1. Регресори у <i>regARIMA</i> моделу	5
3.2. Избор и провера адекватности <i>regARIMA</i> модела	9
4. <i>X-11</i>	18
4.1. <i>B</i> фаза: Прелиминарна оцена пондера за екстремне вредности	19
4.2. <i>C</i> фаза: Финална оцена екстремних вредности.....	31
4.3. <i>D</i> фаза: Финална оцена компоненти	34
5. Провера адекватности десезонирања	42
5.1. Дијагностички тестови – <i>M</i> статистике	43
5.2. Анализа спектралних функција.....	44
5.3. Ревизије.....	45
6. Закључак	49
Библиографија	51

1. Увод

Сезонски карактер временских серија оставља бројне импликације на њихову анализу и, барем у кратком року, отежава њихово праћење. На пример, податак да је индустријска производња у јануару била за 25% нижа у односу на децембарску говори нам, сам за себе, врло мало о тренутном стању у овом сектору и свакако не указује на то да је наступила оштра економска криза, исто као што податак да је температура од августа до децембра била у константном паду не значи да се крећемо ка леденом добу. Наиме, како је индустријска производња у јануару по правилу знатно нижа него у децембру, њихово директно поређење на бази сирових података нема много смисла.

Зато се у пракси најчешће прибегава поређењу са истим периодом претходне године („међугодишњи раст“). Како исти период у две узастопне године има приближно исти сезонски фактор, на овај начин он се елиминише, па подаци постају упоредиви. На пример, податак да је индустријска производња у јануару текуће године за 5% виша у односу на исти месец претходне године, нам свакако више говори од податка из претходног пасуса. Ипак, и овај приступ има извесне мањакости, пре свега, зато што тако добијамо слику о кретању показатеља у последњих годину дана, а не само у последњих неколико месеци, што нас највише занима, па се на овај начин лишавамо анализе текуће економске активности. Постизање упоредивости различитих месеци у току године, постиже се једино искључивањем редовних сезонских осцилација из серије, односно њеним *десезонирањем*.

С десезонирањем временских серија методом *X-12-ARIMA* економисти се редовно сусрећу. Међутим, већина њих, укључујући и аутора овог рада, овај метод готово по правилу користи само притиском на дугме у неком од софтверских пакета. Притом, кораци метода између оригиналне и десезониране серије, у којима се могу уважити специфичности серије, остају у црној кутији. Насупрот томе, у овом раду десезонирамо серију индустријске производње Србије приказујући сваку фазу метода, укључујући притом све релевантне информације за саму серију. Тиме се може побољшати оцена десезониране серије и осталих њених компоненти, а самим тим економско закључивање учинити поузданијим.

Метод *X-12-ARIMA* развио се као надоградња претходно коришћених, *X-11* и *X-11-ARIMA* метода. Метод *X-11* развијен је од стране *U.S. Bureau of the Census* (Shiskin, Young and Musgrave, 1967) и представљао је, као и његови претходници (*X-0* – *X-10*) највећим делом низ пондерисаних покретних просека. На крају серија, у недостатку будућих и прошлих вредности, коришћени су асиметрични покретни просеци, што се узимало као највећа мана овог метода. Каснији методи, *X-11-ARIMA* и *X-11-ARIMA/88* развијени од стране *Statistics Canada* (Dagum, 1980, 1988, 2000), исправили су овај недостатак, “прогнозирајући” прошле и будуће вредности на основу оцењеног *regARIMA* модела, чиме је омогућена употреба симетричних покретних просека. Метод *X-12-ARIMA*, развијен од стране *U.S. Bureau of the Census* (Findley, Monsell, Bell and Otto, 1990) је у односу на *X-11-ARIMA* смањио осетљивост *regARIMA*

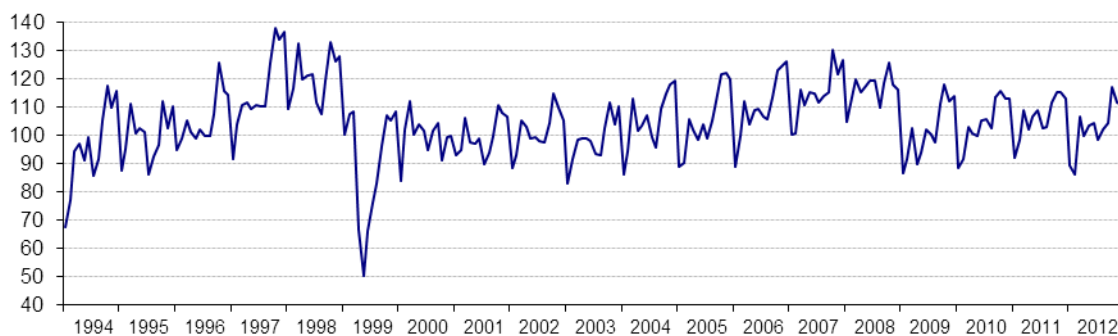
дела и уопште финалних десезонираних серија на искоке у серијама, а и увео је опцију дефинисања регресора од стране корисника.

Поред горепомнутих радова, детаљни прикази X-12-ARIMA метода су углавном такође објављивани од стране статистичких завода. Најдетаљније објашњење метода, корак по корак, дато је у приручницима *U.S. Bureau of the Census – X-12-ARIMA reference manual* (најновија верзија је из 2011.). Ово је приручник пре свега намењен за употребу софтвера X-12-ARIMA (који смо ми користили за овај рад), који првенствено служи за објашњење кодова у програму, али је уједно и веома корисно средство у разумевању самог метода. Laidray и Quenneville (1999) детаљно пролазе кроз сваки корак X-11 метода¹ на примеру једне месечне временске серије. Овај рад, који је илустрован конкретним бројкама, је по нашем мишљењу најилустративнији за разумевање X-11 фазе. *National statistics (UK)* (2007) објавио је такође веома детаљан приказ X-12-ARIMA са јасно илустрованим фазама метода.

У поређењу с наведеним радовима, ми се у приказу X-12-ARIMA метода у знатно већој мери базирамо на графичким приказима фаза метода. Серија индустријске производње Србије изабрана је као погодна за илустрацију метода X-12-ARIMA због изражене сезоналности у њеном кретању условљене пре свега сезоном годишњих одмора, значајног утицаја празника на њено кретање, као и због неколико упадљивих искока (нажалост по правилу негативних), чије моделирање представља значајан елемент овог метода, како у фази предприлагођавања, тако и у фази X-11.

Значајне ирегуларности у серији индустријске производње Србије, чине ову серију нарочито илустративном за опис фазе пред-предприлагођавања (*regARIMA*). Као што ћемо видети, *regARIMA* модел је сам по себи веома погодан за статистичку оцену утицаја ванредних догађаја (нпр. НАТО бомбардовање 1999, петооктобарске политичке промене 2000, почетак светске економске кризе, талас захлађења у фебруару 2012. итд) на кретање временске серије.

Слика 1.1. Серија индустријске производње Србије



¹ X-11 фаза метода X-12-ARIMA веома је слична X-11 методу који је раније доминантно био коришћен за десезонирање.

Мада Републички завод за статистику користи веома софистициран приступ при десезонирању индустријске производње и осталих временских серија, до сада није било објављиваних радова на тему метода *X-12-ARIMA* који би на нивоу детаљности који ми користимо у овом раду анализирали десезонирање неке српске временске серије.

Структура рада прати структуру самог метода *X-12-ARIMA*. Након кратког описа метода на почетку, у другом делу је на примеру детаљно показана фаза прилагођавања, начин на који смо укључили регресоре у *regARIMA* модел, аутоматска оцена искока, уз графички приказ оцењених искока и прилагођене серије која улази у фазу *X-11*, као и провера адекватности *regARIMA* модела. У трећем делу детаљно пролазимо кроз кораке *X-11* фазе у којој се: наизменично и више пута оцењују компоненте серије (трендови, сезонска, ирегуларна), оцењују и замењују екстремне вредности код сезонско-ирегуларне и ирегуларне компоненте, примењују тестови сезоналности. На крају, у четвртном делу проверавамо адекватност десезонирања серије, кроз: *M*-статистике, спектралну анализу и ревизије.

2. Сврха и фазе процеса *X-12-ARIMA* метода

Метод *X-12-ARIMA* декомпонује временску серију на: тренд-циклус (*T*), сезонску (*S*) и ирегуларну (*I*) компоненту. У пракси се користе три типа декомпозиције:

мултипликативни: $Y_t = T_t \times S_t \times I_t$;

адитивни (и лог-адитивни): $Y_t = T_t + S_t + I_t$;

псеудо-адитивни: $Y_t = T_t \times (S_t + I_t - 1)$.

Први тип користи се код највећег дела економских серија, с обзиром да описује кретање серија чије се сезонске осцилације повећавају са њиховим нивоом, што је код економских серија најчешћи случај. Десезонирану серију, у том случају, добијамо дељењем оригиналне серије Y_t сезонским фактором S_t .

Адитивна и лог-адитивна декомпозиција базирају се на сабирању три компоненте, с том разликом што се код лог-адитивне најпре изврши логаритамска трансформација оригиналне серије. Псеудо-адитивна декомпозиција користи се код серија које у одређеним месецима сваке године имају мале или чак нулте вредности. С обзиром на ниске вредности сезонских фактора, мултипликативна декомпозиција код ових серија може довести до изразито високих вредности десезониране серије Y_t/S_t .

Поред тренд-циклус, сезонске и ирегуларне компоненте, које су стохастичког карактера, на понашање већине економских серија утичу и детерминистички фактори, као што су дужина месеца, број радних дана и празници. На пример, већа индустријска

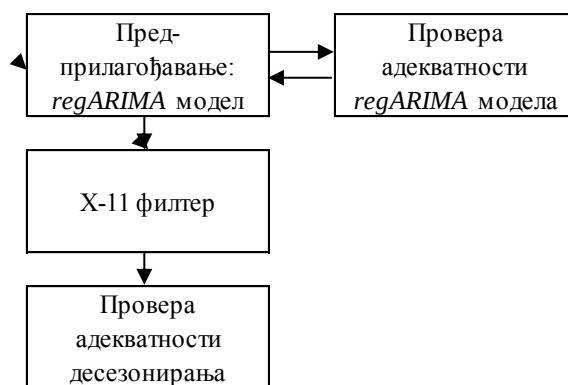
производња у јануару текуће у односу на јануар претходне године може бити последица већег броја радних дана или чињенице да су прошле године, празници, као што су Нова година и Божић пали на дане викенда за разлику од текуће године. Ако ове факторе обележимо са K_t , мултипликативна декомпозиција постаје:

$$Y_t = T_t \times S_t \times K_t \times I_t; \quad (2.1)$$

где K_t обухвата календарске ефекте, као што су број радних дана, дужина месеца, ефекат празника.

Сам поступак десезонирања састоји се из два основна дела: *regARIMA* и *X-11*. Поред тога, *X-12* садржи и бројне тестове за накнадну проверу адекватности *regARIMA* модела и целог десезонирања. Уколико ови тестови у некој фази покажу неадекватност поступка, поступак пред-прилагођавања или целог десезонирања се понавља са измењеним опцијама.

Слика 2.1. Фазе поступка X-12-ARIMA метода



3. Предприлагођавање: *regARIMA* модел

Прва фаза *X-12* метода се назива *пред-прилагођавање (pre-adjustment)*, зато што за циљ има да, пре уласка у *X-11*, серију “прочисти” од искока, календарских ефеката празника и радних дана, и других детерминистичких компоненти, као и да дâ прогнозу (унапред и уназад) овако пречишћене серије. У ту сврху користимо *regARIMA* модел.

RegARIMA модел је комбинација обичне регресије:

$$Y_t = \sum_{i=1}^r \beta_i X_{it} + z_t \quad (3.1)$$

и сезонског *ARIMA* модела резидуала из регресије:

$$\phi_p(L)\phi_p(L^S)(1-L)^d(1-L^S)^D z_t = \theta_q(L)\theta_q(L^S)\varepsilon_t.$$

Ова два израза заједно дефинишу општи облик *regARIMA* модела:

$$\phi_p(L)\phi_p(L^S)(1-L)^d(1-L^S)^D \left[Y_t - \sum \beta_i X_{it} \right] = \theta_q(L)\theta_q(L^S)\varepsilon_t. \quad (3.2)$$

У загради су резидуали z_m из регресије (3.1), с тим што треба имати у виду да се коефицијенти из регресије и *ARIMA* модела оцењују заједнички (симултано). Облик (3.2) сугерише да је *regARIMA* генерализација сезонског *ARIMA* модела.

RegARIMA модел можемо да посматрамо и као генерализацију обичне регресије са грешком која следи стационарни сезонски *ARIMA* процес, ако (3.2) преуредимо на следећи начин:

$$(1-L)^d(1-L^S)^D Y_t = \sum \beta_i (1-L)^d(1-L^S)^D X_{it} + w_t \quad (3.3)$$

при чему је:

$$\phi_p(L)\phi_p(L^S)w_t = \theta_q(L)\theta_q(L^S)\varepsilon_t.$$

Пред-прилагођавање подразумева избор регресора, избор одговарајућег типа *ARIMA* модела, оцену коефицијената уз регресоре и из *ARIMA* модела, одузимање регресионих ефеката и, на крају, прогнозу прочишћене серије изван узорка.

3.1. Регресори у *regARIMA* моделу

Регресори који се укључују у *regARIMA* модел могу се поделити у више група: детерминистичке променљиве, календарски ефекти, ефекти празника и искоци.

Детерминистичке променљиве

Као регресори у *regARIMA* моделу могу се јавити уобичајене детерминистичке компоненте, као што су константа, тренд или сезонске вештачке променљиве. Уколико *ARIMA* део модела не укључује диференцирање, у модел се укључује обична константа. У супротном, као регресор треба укључити *тренд-константу*, која се формира тако да након диференцирања у складу са *ARIMA* моделом, има све вредности једнаке јединици:

$$(1-L)^d(1-L^S)^D \text{ тренд-константа} = 1$$

На пример, у случају кварталних серија, за $d=1$ и $D=0$, ова променљива узима вредности: 1,2,3,4,5,...; за $d=0$ и $D=1$ вредности: 1,1,1,1,2,2,2,2,3,3,3,3,...; док за $d=1$ и $D=1$, *тренд-константа* узима вредности 5,6,7,8,14,16,18,20,27,30,33,36,....

На крају, у регресију можемо укључити сезонске вештачке променљиве, које су овде нешто другачије дефинисане него што је уобичајено. Оне узимају вредност 1 у одговарајућем месецу, вредност -1 у децембру и 0 у осталим месецима. Треба имати у виду, да ако модел укључује сезонско диференцирање, ове променљиве не треба укључивати у модел, зато што након сезонског диференцирања узимају нулту вредност у сваком периоду.

Календарски ефекти

Календарски ефекти могу значајно утицати на вредности серије, па би њихово не-укључивање у регресију водило погрешној оцени *ARIMA* модела и касније (у X-11 фази) неоправданом приписивању ових ефеката сезонској или ирегуларној компоненти. Календарски регресори који се могу користити су: број дана, број радних дана, број појединачних дана у недељи, празници.

Дужина месеца (квартала) код појединих серија може бити веома битна. На пример, производња електричне енергије је у просеку већа у марту него у фебруару, што, на први поглед, није логично с обзиром да она у највећој мери зависи од температуре, која је нижа у фебруару. До ове нелогичности долази управо због чињенице да март има за 10% више дана него (не-преступни) фебруар. Тек када бисмо серију кориговали тако да искључимо ефекат броја радних дана, добили бисмо праву слику. Променљива *дужина месеца* (*length of month*) узима вредности $m_t - \bar{m}$, где је m_t дужина месеца t (у данима), а $\bar{m}$ просечна дужина месеца (30,4375 дана). Програма X-12 користи центриране календарске регресоре, односно средина регресора је једнака нули, како би се избегло да серија прочишћена од календарских ефеката промени општи ниво.

Ипак, код многих серија, као што је индустријска производња, релевантнији је *број радних дана*, зато што се викендом углавном не производи. Број радних дана у месецу може се кретати од 20 до 23, па по овом основу вредности серије могу знатно варирати. Овде се разлике не јављају само између различитих месеци, већ се могу јавити и за исти месец у различитим годинама. На пример, новембар 2004. имао је 22 радна дана, а 2003. године 20, па се, свакако, део међугодишњег раста производње од 12% у новембру 2004. (у односу на исти месец претходне године) може приписати календарском ефекту. Радне дане убацујемо у регресију у следећој форми

$$(\text{број радних дана}) - 5/2(\text{број субота и недеља}). \quad (3.4)$$

На овај начин обезбеђује се да средина регресора буде једнака нули. Приметимо да се у случају фебруара може јавити нелогичност. Уколико фебруар у преступној години

има, на пример пет недеља (и по четири остала дана), онда ће коефицијент рачунат преко (3.4) бити нижи него у фебруару непреступне године, што је нелогично с обзиром да један дан више, макар то била и недеља, не може негативно да утиче на производњу. Зато је уз овај регресор пожељно у регресију укључити и вештачку променљиву за преступну годину, која узима вредност 0,75 у фебруару преступне и -0,25 у фебруару не-преступне године.

Претпоставка када користимо овај регресор је да је ефекат појединачних радних дана у недељи исти, као и да је исти ефекат суботе и недеље. Уколико, међутим, сматрамо да се економска активност разликује по данима, можемо користити посебне регресоре за сваки дан у седмици. Број појединачних дана у месецу може износити 4 или 5. Једна субота или недеља више за собом повлачи један радни дан мање, па стога коефицијенти уз одређене дане могу бити негативни. Једна опција била би да користимо седам вештачких променљивих за сваки дан у седмици. Препоручује се, међутим, употреба шест "контраст" променљивих:

број понедељака – број недеља,, број субота – број недеља,

у комбинацији са седмим регресором – дужином месеца или њеном десезонираном верзијом – регресором за преступну годину. Алтернатива увођењу седмог регресора је да се вредности за фебруар коригују множењем са 28,25/28 у случају не-преступне године, односно са 28,25/29 у случају преступне године.

Ефекат празника

Број нерадних дана у току месеца није дефинисан само бројем дана викенда, већ и бројем нерадних дана због празника. У Србији тренутно постоји шест нерадних празника: Нова година (1. јануар), Божић (7. јануар), Дан уставности (15. фебруар), Ускрс (промењиво), Дан рада (1. мај) и Дан примирја у Првом светском рату (11. новембар). Ускрс је једини "покретан" празник, тј. једини који се не слави увек истог датума. Овај празник може пасти у март, април или мај, с тим што је највећа вероватноћа да падне у април. Оно што је релевантно кад су празници у питању је број нерадних дана повезаних са одређеним празником. Како Ускрс увек пада у недељу, нерадни дани су петак и понедељак око Ускрса. За Нову годину се по правилу не ради прва два радна дана у јануару. Овде треба имати у виду и ефекат "повезивања" празника са викендом, што је код нас устаљена пракса. Уколико, на пример, 1. јануар пада на уторак, многи запослени ће узети слободан и понедељак, 31. децембар, док ће многе фирме и саме одредити овај дан као нерадан. За разлику од осталих празника, уколико Божић падне на викенд, први дан после викенда није нерадан.

Код формирања празничних регресора за мало дуже серије, треба имати у виду значајне измене распореда празника, до којих је долазило у скоријој прошлости: Дан

републике (29. новембар) и Дан борца славили су се до 2000, Дан уставности уведен је 2002, религиозни празници, Божић и Ускрс, „славе“ се од 2001. док је Дан примирја у Првом светском рату нерадан дан од 2012. године. Пре десетак година, славило се само два од шест тренутно постојећих празника.

Празнични регресор можемо да формирамо као број нерадних дана у месецу, који су повезани са одређеним празником, при чему се дани викенда не рачунају да би се избегло дуплирање нерадних дана са регресором који дефинише број радних дана. Тако ће нпр. регресор Дан уставности узети вредност 1 у фебруару 2002, када је овај празник пао на радни дан, а вредност 0 у фебруару 2003, када је пао у суботу.

Искоци

Искоци (outliers) представљају значајнија одступања временске серије од уобичајеног кретања. Ово питање је код нас веома битно, зато што већина, ако не и све економске серије имају барем неколико структурних ломова и екстремних искакања у појединим периодима.

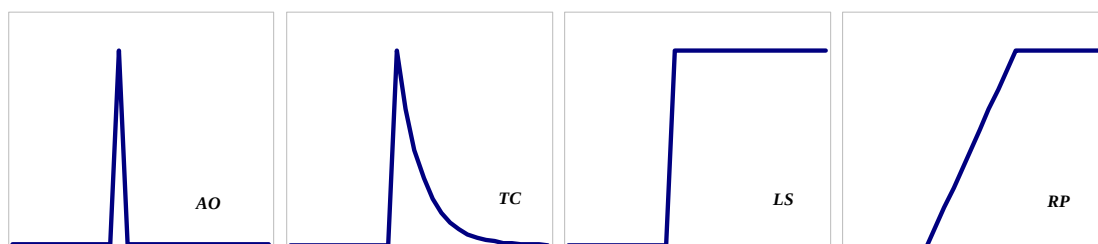
Искоци могу бити привремени или трајни. Метод X-12 разликује четири основна типа искока: *адитивни искок (AO – additive outlier)*, *привремена промена (TC – temporary change)*, *промена нивоа (LS – level shift)* и *рампа (RP – ramp)*. Вештачке променљиве које се користе за ове искоке, дефинишу се на следећи начин²:

$$AO_t^{(t_0)} = \begin{cases} 1 & \text{zat} = t_0 \\ 0 & \text{zat} \neq t_0 \end{cases} \quad LS_t^{(t_0)} = \begin{cases} -1 & \text{zat} < t_0 \\ 0 & \text{zat} \geq t_0 \end{cases}$$

$$TC_t^{(t_0)} = \begin{cases} 0 & \text{zat} < t_0 \\ \alpha^{t-t_0} & \text{zat} \geq t_0 \end{cases} \quad RP_t^{(t_0, t_1)} = \begin{cases} -1 & \text{zat} \leq t_0 \\ (t-t_0)/(t_1-t_0)-1 & \text{zat}_0 < t < t_1 \\ 0 & \text{zat} \geq t_1 \end{cases}$$

Код адитивног искока, промена се дешава само у једном периоду (t_0) и већ у наредном серија се враћа у нормалу. Регресор AO је импулсна вештачка променљива, која узима вредност 1 у периоду шока (t_0) и вредност 0 у свим осталим месецима. У случају привремене промене (TC), након почетног шока серија се враћа у равнотежу али не одмах као у случају AO, већ постепено, по одређеној стопи, α , која узима вредности између 0 и 1. Програм X-12 по аутоматизму узима стопу $\alpha=0,7$ али је могуће дефинисати и друге стопе.

² U.S. Census Bureau. 2011., “X-12-ARIMA Reference Manual”.

Слика 3.1: Типови скока у *regARIMA* моделу

Код промене нивоа (*LS*), ефекат шока је трајан – вредности серије након шока трајно мењају ниво, на горе или доле. Регресор *LS*, који одговара овом типу структурног лома, узима вредности -1 до периода промене (t_0) и 0 након тога. Приметимо да овај регресор узима вредности -1 и 0, а не 0 и 1, како је уобичајено, да би општи ниво средине регресионе функције био конзистентан са текућим нивоом серије³. На крају, *рампа* (*RP*) такође представља трајну промену нивоа серије, али не у једном месецу, као код претходног случаја, већ постепено. Овај регресор узима вредности -1 до почетка промене (t_0), 0 након трајне промене нивоа (t_1), а вредности између -1 и 0 у прелазном периоду.

Као регресоре у *regARIMA* могуће је унапред задати тип и тренутак настанка структурног лома – на пример, *LS* за април 1999. године, у ком случају би овај регресор узео вредности -1 до марта 1999, а 0 од априла 1999. па надаље. Поред тога, програму можемо задати да аутоматски испита све могуће комбинације типова и тренутака настанка структурног лома (за типове *AO*, *TC* и *LS*) у анализираном периоду, док је за *RP* у сваком случају неопходно специфицирати период лома.

3.2. Избор и провера адекватности *regARIMA* модела

Избор ARIMA модела

Програму *X-12* могуће је унапред задати одређен број *ARIMA*(p,d,q)(P,D,Q) модела, између којих би изабрао најбољи на основу одређених критеријума. Програм *X-12* по аутоматизму испитује следећих пет сезонских *ARIMA* модела:

$$(0\ 1\ 1)(0\ 1\ 1)$$

$$(0\ 1\ 2)(0\ 1\ 1)$$

$$(2\ 1\ 0)(0\ 1\ 1)$$

$$(0\ 2\ 2)(0\ 1\ 1)$$

$$(2\ 1\ 2)(0\ 1\ 1)$$

³ U.S. Census Bureau. 2011., “X-12-ARIMA Reference Manual”.

с тим што можемо и сами дефинисати модел који сматрамо оговарајућим. Како модели могу имати различите редове диференцирања, информационе критеријуме не можемо применити при избору модела. Уместо тога, дефинисани су критеријуми адекватности које модел мора да задовољи да би ушао у „ужи избор“.

Модел је дефинисан као адекватан уколико задовољава три услова:

1) просечна апсолутна процентуална грешка екстраполираних вредности у последње три године је мања од 15%

2) *Box-Ljung*-ова Q -статистика на 24. доцњи није статистички значајна, тј. t -вредност израчунате статистике је већа од 5%

3) нема знакова прекомерног диференцирања – сума оцењених несезонских MA параметара мања је од 0,9.

Како све критеријуме може задовољити више модела, између њих се може изабрати модел са најмањом унутар-узорачком грешком прогнозе или, као што смо ми урадили у овом папиру, први модел са листе који задовољава сва три критеријума.

Тестирање значајности параметара

Укључивање искока у регресију базира се на t -статистици. Програму се може задати аутоматско тестирање за AO , TC и LS типове искока, за цео период T . Овај поступак подразумева оцену коефицијента и израчунавање t -статистике за сва три типа екстремне вредности за сваки месец (квартал) у посматраном интервалу. Како се у том случају тестира велики број променљивих⁴, није оправдано користити уобичајену критичну вредности од 1,96 за ниво значајности 5%, па програм користи знатно више критичне вредности, 3,9, што одговара нивоу значајности од 0,01%.

Провера адекватности $regARIMA$ модела

Провера адекватности укључује анализу резидуала w_t из (2.3), са циљем да се провери да ли они следе процес бели шум, што подразумева нормалност расподеле са нултом средином и одсуство аутокорељације. Нормалност резидуала w_t се тестира *Geary-јевом* статистиком и коефицијентом спљоштености. *Geary-јева* статистика рачуна се као⁵

⁴ За сваки период, њих T , тестирају се три типа искока, што је укупно $3T$ или 684 у примеру серије коју ћемо анализирати касније. При нивоу значајности од 5%, могли бисмо открити око 34 статистички значајних искока, чак иако серија у стварности не поседује ниједан.

⁵ U.S. Census Bureau. 2011., “X-12-ARIMA Reference Manual”.

$$a = \frac{\frac{1}{T} \sum_{i=1}^T |w_i - \bar{w}|}{\sqrt{\frac{1}{T} \sum_{i=1}^T (w_i - \bar{w})^2}}$$

где је $\bar{w}$ просечна вредност резидуала w_i , при чему је нулта хипотеза овог теста да је расподела нормална. Вредност коефицијента спољоштености креће се око три ако је расподела нормална. Аутокорелисаност резидуала проверава се анализом *ACF* и *PACF* корелограма резидуала са *Box-Ljung*-овом *Q*-статистиком.

Поред тога, проверава се спектар резидуала, са циљем да се испита да ли су елиминисане сезонске и тзв. *фреквенције радних дана*. Типична серија са сезонским осцилацијама има спектралну функцију са локалним максимумима на неким или свим сезонским фреквенцијама (број циклуса по периоду), а код месечних серија то су: 1/12, 2/12, ..., 6/12.

Поред сезоне, унутаргодишње осцилације серије могу бити и резултат различитог броја радних дана по месецима. У том случају, серија ће имати спектралне врхове и на фреквенцијама радних дана, а то су код месечних серија 0,348 и 0,432. Прва фреквенција (0,348) израчуната је на основу броја седмица у месецу просечне дужине (30,4375 дана), тј. 4,348 (=30,4375/7). Овај број заправо представља број циклуса који седмице, односно радни дани направе у току месеца просечне дужине. Како радимо са месечним серијама, не занимају нас осцилације унутар месеца, већ унутар године. Због тога, елимишемо онај део фреквенције који се односи на унутармесечне осцилације (4) и посматрамо само његов децимални део (0,348), зато што он одређује унутаргодишње осцилације. Другу фреквенцију радних дана (0,432) израчунали су *Cleveland* и *Devlin* (1980), као фреквенцију на којој ће серија тока (нпр. индустријска производња) највероватније имати спектрални врх, ако има значајну компоненту радних дана. Ова фреквенција није релевантна за серије стања.

Корекција и прогноза серије

Пошто је модел изабран, а коефицијенти уз регресоре и *ARIMA* модела оцењени, као крајњи резултат фазе пред-прилагођавања, оригинална серија коригује се за регресионе ефекте и прогнозирају се вредности изван узорка, унапред и уназад. Од оригиналне серије најпре одузимамо регресионе ефекте на основу оцењених коефицијената из регресије:

$$Y_t^{B1} = Y_t - \sum_i \beta_i X_{it}.$$

Једини регресиони ефекат који се не одузима је тренд-константа зато што би се на тај начин променио ниво серије. Затим, на основу оцењеног *ARIMA* модела, прогнозирамо унапред и уназад (*forecast* и *backcast*) кориговану серију (Y_t^{B1}). У случају серија дужих од 15 година, ради се само прогноза унапред. Нова, коригована серија улази у следећу фазу, *X11*, као почетна серија.

У наставку рада ћемо приказати цео процес метода X-12-ARIMA на примеру серије индустријске производње. Најпре ћемо објаснити примену *RegARIMA* модела на овој серији.

Примена RegARIMA на серији индустријске производње Србије

Дефинисање регресора

Први корак у оцени *RegARIMA* модела је избор променљивих које се укључују у регресију. Ми смо укључили две групе регресора – једну која би обухватила календарске ефекте, а другу која би обухватила искоке.

Када су у питању календарски ефекти, већ смо рекли да је у случају индустријске производње број радних дана релевантнији од дужине месеца. Број радних дана у току месеца зависи од броја понедељка-петка у току месеца (којих може бити 20-23, као што смо рекли) и од броја нерадних дана због празника. Због тога смо укључили две променљиве: број радних дана у току месеца (променљива *RD*) и број нерадних дана због празника (*PR*). Променљива *RD*, дефинисана је као у (3.4). Променљива *PR* иницијално је дефинисана је као број дана унутар радне седмице (понедељак-петак) којима се не ради због празника. Да би се избегло дуплирање, ако празник пада на викенд (и не помера се за после викенда) онда се он не рачуна као нерадни дан. На пример, ако Божић пада у среду, променљива *PR* узима вредност један, а ако пада у недељу, узима вредност нула, зато што се недеља иначе води као нерадни дан у оквири променљиве *RD*.

Поред тога, ова променљива укључује и споредне ефекте празника. На пример, ако Први мај пада у среду, нерадни дани су среда и четвртак, али, поред тога, многи запослени узимају и петак као слободни дан, ради тзв. спајања са викендом. У овом случају, за среду и четвртак, променљива *PR* узима вредност 1, а за петак вредност 0,3, па променљива *PR* за цео мај те године узима вредност 2,3. Поред тога, познато је да 31. децембра, иако није празнични дан, већина фирми ради скраћеним радним временом уз прославе у току радног времена, тако да за децембар (уколико 31. не пада на викенд) стављамо вредност 0,5⁶.

Имали смо у виду промене празничног календара, до којих је дошло након 2000. године. Тако су, на пример, до 2000. године у променљиву *PR* укључени ефекти Дана републике (29. новембар), а искључени ефекти Дана уставности (15. фебруар). На крају, да би се средина променљиве свела на нулу, променљива *PR* трансформисана је тако што је подељена сопственом средином.

Када су у питању искоци, на основу графичког приказа, најочигледнији је драстичан пад индустријске производње 1999. године током бомбардовања. Овај

⁶ Коефицијенти за споредне ефекте празника су изабрани произвољно.

искок, на први поглед, највише личи на привремену промену (TC), с тим што је упадљив и пад нивоа серије (LS) по завршетку бомбардовања. Ипак, за разлику од типичног TC искока, код кога серија бележи значајан пад у једном месецу, а затим се постепено враћа у нормалу, у овом случају, пад се иницијално дешавао у три месеца (март, април, мај), тако да избор вештачких променљивих које би обухватиле овај искок серије није очигледан.

Још један познати искок је крај 2000. године, када је након петооктобарских политичких промена дошло до привременог успоравања индустријске производње. У овом случају, уместо уобичајеног сезонског пораста крајем године, серија је у овом периоду стагнирала. Ни за овај искок није очигледан избор вештачке променљиве. Због тога смо се определили да вештачке променљиве за искоке не задајемо *a priori*, већ да програм сам оцени све комбинације искока и изабере статистички значајне.

Стопу α искока TC смо подесили на 0,55, уместо уобичајених 0,7. За ову стопу смо се определили након неколико итерација целог процеса десезонирања, на основу графичке анализе и статистичке значајности коефицијената уз TC променљиве. Уобичајена стопа од 0,7 не описује добро ефекат бомбардовања у 1999. години, зато што представља преспоро одумирање шока у односу на оно што се стварно десило. У случају ефекта политичких промена 2000. била би неопходна још нижа стопа. У том смислу, немогућност задавања различитих стопа α за различите TC искоке недостатак је програма $X-12$, по нашем мишљењу.

Последњи упадљив и трајан искок повезан је са почетком светске економске кризе, крајем 2008. и почетком 2009. Пад индустријске производње у овом периоду био је снажан и трајао је десетак месеци, пре него што је у другом тромесечју 2009. године започет умерени опоравак са (трајно) нижег нивоа. Стога смо у $RegARIMA$ укључили регресор рампу ($RP^{avg08-maj09}$) за период август 2008. – мај 2009. Подсећамо да се овај тип искока не оцењује аутоматски, већ га неопходно посебно дефинисати.

Избор $ARIMA$ модела

При избору $ARIMA$ модел примењена је методологија избора првог модела са листе (стр. 16) који задовољава сва три критеријума адекватности. Испоставило се да сва три критеријума адекватности задовољава већ први модел на листи – $(0\ 1\ 1)(0\ 1\ 1)$.

Табела 3.1. Критеријуми адекватности оцењеног $SARIMA$ модела $(0\ 1\ 1)(0\ 1\ 1)$

Апсолутна просечна процентуална грашка екстраполираних вредности у последње три године	p -вредност израчунате $Box-Ljung$ -ове Q -статистике	Оцењене вредности коефицијената $SARIMA$ модела
2,70%	27,4%	$\theta_1 = 0,405$, $\Theta_1 = 0,813$

Резултати регресије

Резултати регресије показују статистичку значајност две календарске компоненте и рампе које смо укључити у регресију, док је аутоматска детектовано још шест статистички значајних искока (табела 3.2).

Табела 3.2. Оцењени коефицијенти *regARIMA* модела

Променљива	Оцењени коефицијент	t-статистика
Регресија		
<i>RD</i>	0.0043	8,47
<i>PR</i>	-0.0154	-4.49
<i>TC^{jan94}</i>	-0.2019	-4.61
<i>LS^{mar99}</i>	-0.1415	-5.04
<i>AO^{apr99}</i>	-0.4177	-14.82
<i>TC^{maj99}</i>	-0.6734	-22.07
<i>TC^{okt00}</i>	-0.2378	-8.55
<i>AO^{feb12}</i>	-0,1213	-8,96
<i>RP^{avg08-maj09}</i>	-0,0286	-4,76
ARIMA		
Несезонска МА компонента (θ_1)	0,4051	6,68
Сезонска МА компонента (Θ_1)	0,8127	19,63

Како смо променљиве регресирали на логаритмовану серију, оцењене коефицијенте можемо тумачити као апроксимацију коефицијената еластичности. У случају променљиве *PR*, тумачење је једноставно – један празнични дан више значи 1,54% производње мање. С обзиром да месец у просеку садржи 21,75 радних дана (понедељак-петак), уколико би се празницима у потпуности обустављала производња, очекивали бисмо да се оцењени коефицијент креће око 4,5%. Како је оцењени коефицијент мањи, можемо да закључимо да се фабрике празницима не закључавају у потпуности или се мањак производње за време празника надокнади осталим данима у месецу.

Тумачење оцењеног коефицијента уз променљиву за радне дане (*RD*), нешто је сложеније. Ако пођемо од једначине (3.4) можемо лако израчунати да један радни дан више значи да променљива *RD* узима вредност већу за 3,5. Према томе, коефицијент уз ову променљиву неопходно је помножити са 3,5 да бисмо га тумачили као еластичност на промену једног радног дана. Ако то урадимо, добијамо коефицијент еластичности од 1,51%, што је веома блиско коефицијенту уз празничну променљиву.

Ако пођемо од декомпозиције (2.1), у овој фази већ имамо оцењену финалну календарску компоненту K_t , с тим што је неопходна одговарајућа трансформација да би се ефекат прилагодио мултипликативном методу. У ту сврху, вештачке календарске променљиве (RD и PR) помножене одговарајућим оцењеним коефицијентима из $regARIMA$ модела (β_{RD} и β_{PR}), неопходно је анти-логаритмовати:

$$efRD = e^{\beta_{RD} \cdot RD} = e^{0,0043RD} \quad efPR = e^{\beta_{PR} \cdot PR} = e^{-0,0154PR} \quad (3.5)$$

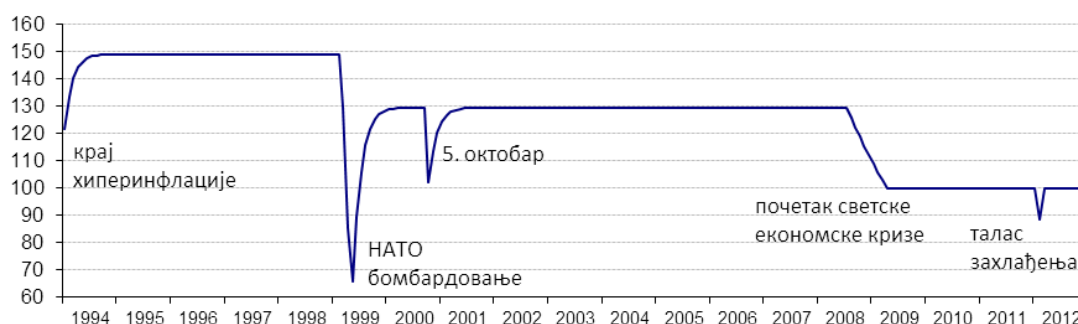
где је e основа природног логаритма, а ef испред променљиве означава да је у њен ефекат. Заједничку календарску компоненту (K_t) добијамо као производ два календарска ефекта:

$$K_t = efRD \cdot efPR, \quad (3.6)$$

Ову компоненту користићемо у следећој фази поступка ($X-11$) приликом финалне декомпозиције.

У $regARIMA$ моделу статистички значајним показало се седам искока, од којих се три односе на НАТО бомбардовање. Први испок, привремена промена TC^{jan94} , налази се на самом почетку серије. Ово се може објаснити тиме да серија почиње у месецу хиперинфлације, која је имала у погубне последице по привреду. Програм монетарне реформе којим је заустављена хиперинфлација, и који је спроведен у фебруару исте године, довео је до постепеног опоравка индустријске производње у тој години. Исти тип искока (TC) оцењен је и за октобар 2000. године, месец петооктобарских политичких промена, које су довеле до привременог пада индустријске производње у том месецу, и постепене нормализације након тога. Рампа (RP) коју смо укључили за период почетка светске економске кризе такође се показала статистички значајном, а оцењени коефицијент показује месечни пад од 2,86%, односно кумулативан од 25,74% у првих девет месеци кризе. На крају, значајним се показао још привремени испок (AO) за фебруар 2012, месец у коме је двонедељни талас захлађења привремено пореметио процес производње.

Слика 3.2. Ефекат оцењених искока из $regARIMA$ модела



Када су у питању ратни месеци, они су оцењени комбинацијом искока. Поред привременог пада који је обухваћен променљивама AO^{apr99} и TC^{maj99} , у току рата, дошло је и до структурног лома у серији, односно до трајног пада на за 14% нижи ниво

($LS^{мар99}$), који се задржао и по завршетку рата. На слици 3.3. заједно су приказани сви оцењени искоци.

Провера адекватности *regARIMA* модела

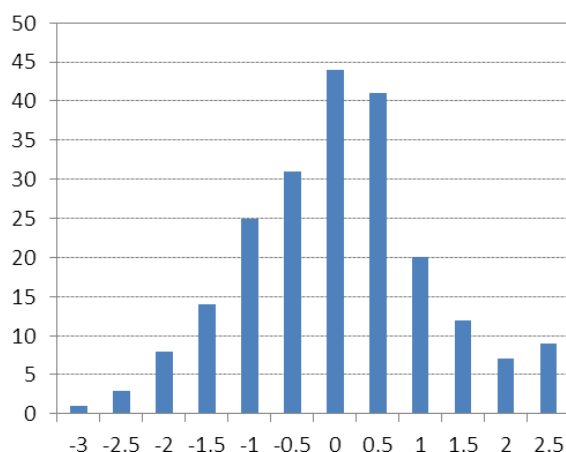
Резултати из табеле (3.2) су оцењени коефицијенти модела (3.2), односно, конкретније, модела $ARIMA(0\ 1\ 1)(0\ 1\ 1)$:

$$(1-L)(1-L^{12})\left(lpY_t - \sum_{i=1}^7 \beta_i X_{it}\right) = (1-\theta_1 L)(1-\theta_1 L^{12})\varepsilon_t \quad (3.5)$$

где је lpY_t логаритмована (l) серија индустријске производње (коригована за ефекте преступне године (p)), X_{it} , $i = 1, \dots, 7$, регресори, а β_i оцењени коефицијенти из табеле (3.2). Као што смо већ рекли, да бисмо прихватили модел, пожељно је да резидуали ε_t из (3.5) задовоље одређене услове: да поседују карактеристике белог шума (нормалност расподеле и одсуство аутокорељације) и да не садрже значајне сезонске или фреквенције радних дана.

Када је у питању нормалност, хистограм стандардизованих резидуала (слика 3.3) указује на могуће одступање њихове расподеле од нормалне. Ипак, оцењене статистике *Geary*-јевог теста нормалности $(0,79)^7$ и узорачки коефицијент спљоштености $(2,99)$ нису довољне да одбацимо нулту хипотезу о нормалности, па закључујемо да су резидуали задовољили овај критеријум.

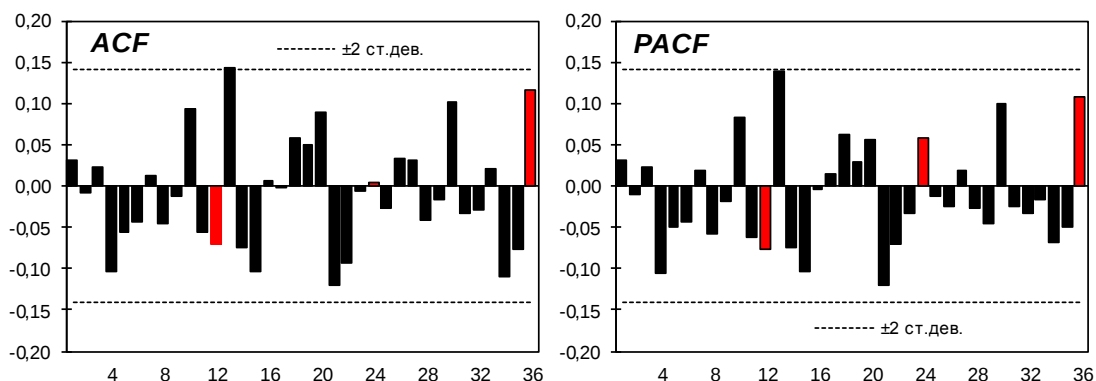
Слика 3.3. Расподела резидуала из *regARIMA* модела



⁷ Програм X-12-ARIMA не приказује критичну вредност *Geary*-јевог теста

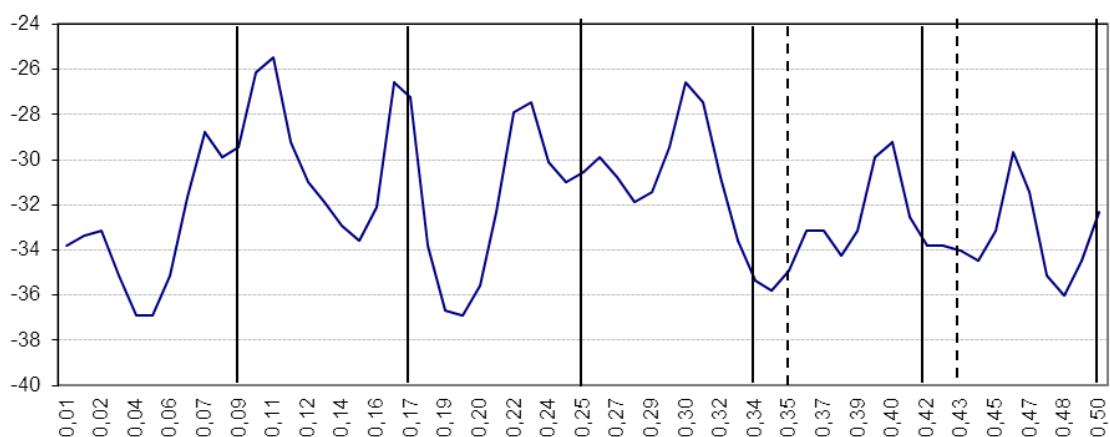
Задовољавајући су и обични и парцијални корелограми резидуала (слика 3.4), који указују на одсуство аутокорејације и на обичним и на сезонским фреквенцијама (црвени стубови).

Слика 3.4. Обична и парцијална аутокорејациона функција резидуала из *regARIMA* модела



Спектрална функција резидуала (слика 3.5) такође показује адекватност регресије, с обзиром да не доминирају сезонске фреквенције.

Слика 3.5. Функција спектралне густине резидуала из *regARIMA* модела



Можемо рећи да резидуали задовољавају основне критеријуме адекватности и да оцењени *regARIMA* модел можемо прихватити као одговарајући.

На крају, на основу резултата овог модела рачунамо прилагођену оригиналну серију, са циљем да је очистимо од ефеката радних дана, празника и искока, и

прогнозирамо унапред две године унапред⁸. Прилагођавање се врши тако што се од оригиналне логаритмоване серије одузимају оцењени ефекти регресора

$$IY_t^{(B1)} = IY_t - \sum_{i=1}^7 \beta_i X_{it}$$

где је IY_t логаритмована (I) оригинална серија, а x_{it} , $i = 1, \dots, 7$, регресори RD , PR и пет искока. Овако прилагођена серија се затим прогнозира две године унапред, на бази оцењеног $ARIMA(0 \ 1 \ 1)(0 \ 1 \ 1)$ модела. На крају, цела серија се анти-логаритмује:

$$Y_t^{B1} = e^{IY_t^{B1}}$$

зато што у следећој фази, $X-11$, примењујемо мултипликативни метод.

Овде се процес пред-прилагођавања завршава. Као коначан резултат добили смо прилагођену серију, прогнозирану две године унапред и уназад. У суперскрипту смо је обележили са $B1$ зато што ова серија улази у почетну, $B1$, фазу процеса $X-11$.

4. X-11

Програм $X-11$ се у највећем делу састоји из низа пондерисаних покретних просека, којима се више пута, прелиминарно, а на крају и финално, наизменично оцењују тренд, сезонска и ирегуларна компонента, као и екстремне вредности серије. $X-11$ се састоји из три фазе, које се обично обележавају као B , C и D фаза⁹. Крајњи резултат прве две фазе (B и C) је оцена екстремних вредности, које се на почетку следеће фазе елиминишу из серије. Оцени екстремних вредности придаје се велики значај у целом $X-12-ARIMA$ методу, па и поред тога што је серија у пред-прилагођавању већ очишћена од највећих искакања, у $X-11$ поново се тестира и, по потреби, елиминише присуство екстремних вредности, које нису оцењене као екстремне у првој фази. У последњој, D фази даје се финална оцена компоненти. Свака од три фазе састоји се из великог броја међу-корака.

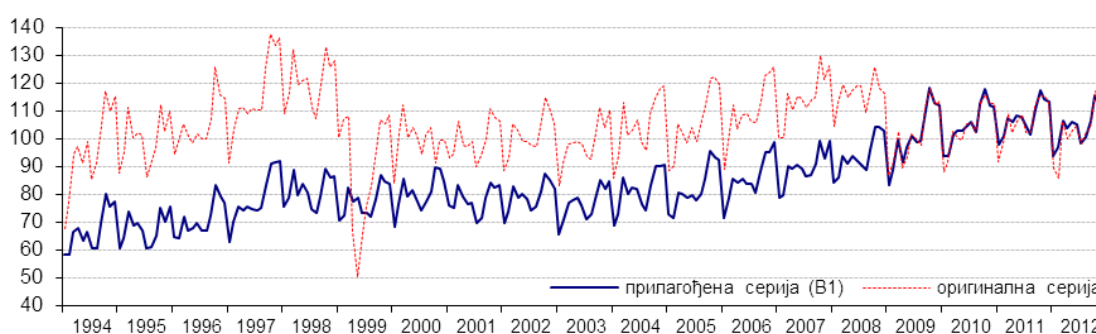
⁸ За серије краће од 15 година ради се екстраполација серије и уназад.

⁹ Са A се обележава фаза пред-прилагођавања.

4.1. В фаза: Прелиминарна оцена пондера за екстремне вредности

В1 Прилагођена оригинална серија

У *X-11* део улази серија пред-прилагођена на основу *regARIMA* модела, очишћена од календарских ефеката и највећих искакања (крај хиперинфлације 1994, НАТО бомбардовање 1999, 5. октобар 2000, почетак светске економске кризе 2008-2009 и талас захлађења у фебруару 2012.), са прогнозама две године унапред.

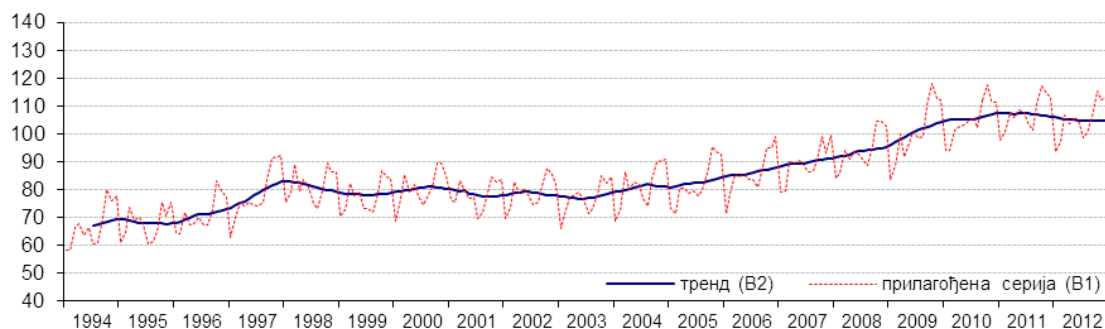


В2 Тренд-циклус – 12-члани централизовани покретни просек

Почетна серија у *X-11* садржи све три основне компоненте (*TSI*). У овој фази циљ нам је да из ње издвојимо тренд компоненту. Прелиминарна оцена тренд компоненте добија се применом 12-чланог центрираног покретног просека на *B1*, који заправо представља обичан двочлани покретни просек, примењен на 12-члани покретни просек, или 2×12 покретни просек:

$$T_t^{B2} = M_{2 \times 12}(y_t^{B1}) = \frac{1}{24} Y_{t-6}^{B1} + \frac{1}{12} Y_{t-5}^{B1} + \dots + \frac{1}{12} Y_t^{B1} + \dots + \frac{1}{12} Y_{t+5}^{B1} + \frac{1}{24} Y_{t+6}^{B1},$$

где је $M_{2 \times 12}$ ознака за 2×12 покретни просек. Како је за ову калкулацију неопходно по шест опсервација с обе стране, на крају серије користе се прогнозиране вредности из *regARIMA* фазе. С обзиром да програм *X-12-ARIMA* за серије дуже од 15 година не ради екстраполацију уназад, за првих шест месеци серије нема опсервација за рачунање овог покретног просека.



На овај начин серија се чисти од сезонских фреквенција. Применом 2×12 покретног просека, из Y_t^{B1} серије смо искључили сезонско-ирегуларну (SI) компоненту.

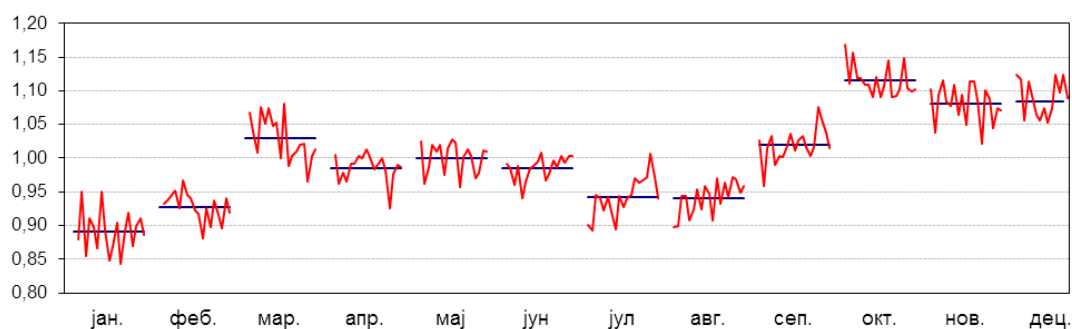
***B3* Сезонско-ирегуларна компонента**

Дељењем серије Y_t^{B1} (TSI) прелиминарном T компонентом из $B2$ добијамо прелиминарну оцену сезонско-ирегуларне (SI) компоненте:

$$SI_t^{B3} = Y_t^{B1} / T_t^{B2}$$

Ако ову серију прикажемо одвојено по месецима (слика 4.1), приметне су значајне сезонске разлике. Као што је за очекивати када је у питању индустријска производња, последње тромесечје у години има највише SI вредности, а летњи месеци и јануар најниже.

Слика 4.1. Сезонско-ирегуларна компонента ($B3$) по месецима



Тест стабилне сезоналности серије

Да бисмо и формално проверили да ли су разлике међу месецима значајне, спроводимо *тест стабилне сезоналности*¹⁰. Ово је тест анализе варијансе где третирамо 12 сезонских узорака сезоне величина $n_1, \dots, n_{12}$. У нашем случају величине сезонских узорака су исте за све сезоне – 19 година. За сваку од тих сезона имамо реализације SI компоненте. На основу њих тестирамо да ли су популацијске средине (SI_i^P) по сезонама различите.

$$\begin{aligned} X_0: SI_1^P &= \dots = SI_{12}^P, && \text{нема стабилног сезонског карактера} \\ X_1: SI_i^P &\neq SI_j^P \text{ за барем један пар } (i, j), && \text{стабилни сезонски карактер.} \end{aligned}$$

Горње хипотезе тестирамо тако што варијансу SI компоненте целог узорка, V^2/n , разлажемо на варијансу од SI између сезона, V_a^2/n , и варијансу унутар сезона тј. резидуалну варијансу, V_r^2/n :

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^{12} \sum_{j=1}^{n_j} (SI_{ij} - \overline{SI})^2 &= \frac{1}{n} \sum_{j=1}^{12} n_j (\overline{SI}_j - \overline{SI})^2 + \frac{1}{n} \sum_{i=1}^{12} \sum_{j=1}^{n_j} (SI_{ij} - \overline{SI}_i)^2 \quad \text{или} \\ \frac{1}{n} V^2 &= \frac{V_a^2 + V_r^2}{n} \end{aligned}$$

Уколико су варијације између различитих сезона релативно значајне у односу на варијације унутар сезона, закључујемо да серија има статистички различите сезонске средине SI , односно да има стабилни сезонски карактер. Ово тестирамо преко F -статистике:

$$F_V = \frac{V_A^2/11}{V_R^2/(n-1)},$$

која има F расподелу са $k-1$ и $n-k$ степени слободе. Уколико је израчуната вредност ове статистике већа од критичне вредности за $k-1$ и $n-k$ степени слободе, одбацујемо нулту хипотезу и закључујемо да серија има стабилну сезоналност. Због могућности да нису испуњене бројне претпоставке за примену анализе варијансе (нпр. одсуство аутокорејације ирегуларне компоненте), потребно је изабрати високу критичну вредност за тестирање, најчешће се узима $\alpha=0,1\%$.

Овде је битно нагласити да прихватање нулте хипотезе не значи да треба да закључимо да серија нема сезону. Прихватање нулте хипотезе може бити последица и изузетно променљиве сезоне која за последицу може имати немогућност статистичког разликовања сезонских средина.

¹⁰ Laidray and Quenneville (1999).

У нашем примеру добили смо следеће резултате:

Табела 4.1. Резултати теста стабилне сезоналности

	Сума квадрата	Степени слободe	F-статистика
Између месеци	9.527,9	11	129,6 (стат. значајно)
Резидуал	1.363,2	204	
Укупно	10.891,2	215	

Чак и на нивоу значајности од 0,1%, можемо да закључимо да индустријска производња има стабилну сезоналност.

В4 Замена екстремних вредности сезонско-ирегуларне компоненте

Следећи корак била би оцена прелиминарне сезонске компоненте (S), филтрирањем SI компоненте покретним просецима. Генерално, проблем са покретним просецима је што могу бити под значајним утицајем ирегуларности у кретању филтрираних серија. Уколико је, на пример, у неком месецу дошло до шока који је изазвао привремени пораст производње, може доћи до неоправданог приписивања једнократног пораста производње сезонским факторима. Зато, да би се избегло да ирегуларности у кретању SI компоненте утичу на оцену сезонског фактора, потребно је у овој фази заменити њихове екстремне вредности.

Поступак идентификовања и замене екстремних вредности обавља се у више корака. Прво оцењујемо S компоненту у две фазе. Најпре, на SI примењујемо 3×3 покретни просек, који је, заправо, трочлани покретни просек примењен на трочлани покретни просек. У питању је сезонски филтер, што значи да се одвојено примењује на различите сезоне. Може се показати да је 3×3 филтер исто што и пондерисани петочлани покретни просек:

$$M_{3 \times 3}^{12}(SI_t^{B3}) = \frac{1}{9}SI_{t-24}^{B3} + \frac{2}{9}SI_{t-12}^{B3} + \frac{3}{9}SI_t^{B3} + \frac{2}{9}SI_{t+12}^{B3} + \frac{1}{9}SI_{t+24}^{B3},$$

где 12 у суперскрипту означава да је у питању сезонски месечни покретни просек. Прелиминарну S компоненту у овој фази добијамо нормализацијом, тако што 3×3 филтер делимо његовим 12-месечним централизованим покретним просеком.

$$S_t^{B4} = M_{3 \times 3}^{12}(SI_t^{B3}) / M_{2 \times 12}[M_{3 \times 3}^{12}(SI_t^{B3})],$$

док ирегуларну компоненту у овој фази рачунамо као:

$$I_t^{B4} = SI_t^{B3} / S_t^{B4}$$

Затим, израчунавамо петогодишњу покретну стандардну девијацију ирегуларне компоненте, с тим што тако израчунату стандардну девијацију приписујемо средњој години из петогодишта:

$$\sigma_\tau^{B4} = \sqrt{\frac{1}{60} \sum_{t=\text{jan}(\tau-2)}^{\text{dec}(\tau+2)} (I_t^{B4} - 1)^2},$$

где је σ_τ^{B4} стандардна девијација за годину τ у фази $B4$. За стандардну девијацију прве две (1994. и 1995.) и последње две године (2011. и 2012.) узимамо да је једнака стандардној девијацији прве следеће године за коју је могуће израчунати петогодишњу стандардну девијацију (1996. и 2010.).

Затим, рачунамо колико поједине ирегуларне вредности одступају од просека у стандардним девијацијама, и на бази одступања им придодајемо пондере¹¹:

- вредностима које од средине одступају мање од 1,5 стандардне девијације приписује се пондер 1 и оне се не коригују;
- вредности које одступају више од 2,5 стандардних девијација од средине (1), сматрају се екстремним и њима се приписује пондер нула;
- вредностима које од средине одступају између 1,5 и 2,5 стандардних девијација приписује се пондер који узима вредности између 0 и 1 у зависности од тога колико ирегуларна компонента одступа од средине (1):

$$p_t^{B4} = (2,5 \cdot \sigma_\tau^{B4} - |I_t^{B4} - 1|) / (2,5 \cdot \sigma_\tau^{B4} - 1,5 \cdot \sigma_\tau^{B4}),$$

где је p_t^{B4} пондер за корекцију екстремних вредности у фази $B4$, при чему месец t припада години τ . У питању је линеарно пресликавање интервала 1,5-2,5 у интервал 0-1. На пример ако ирегуларна компонента одступа од средине 2 стандардне девијације, пондер узима вредност 0,5

¹¹ *Laidray and Quenneville (1999).*

Слика 4.2. Пондери за корекцију SI компоненте у зависности од одступања ирегуларне компоненте од средине



У нашем примеру, на овај начин идентификовано је 30 вредности ирегуларне компоненте које одступају од средине више од 1,5 стандардних девијација. Од тог броја, чак половина се односи на јануар (10) и март (5), док само за фебруар није идентификована ниједна екстремна вредност (слика 4.3).

Вредности SI компоненте, чија ирегуларна компонента одступа више од 1,5 стандардне девијације од један, замењују се пондерисаним покретним просеком саме те вредности, са одговарајућим пондером, и по две претходне и будуће вредности са пуним пондером (1).

$$SI_t^{B4} = V_t^{B4} \cdot SI_t^{B3} + (1 - V_t^{B4}) \cdot (SI_{t-24}^{B3} + SI_{t-12}^{B3} + p_t^{B4} SI_t^{B3} + SI_{t+12}^{B3} + SI_{t+24}^{B3}) / (4 + p_t^{B4}), \quad (4.1)$$

где је V_t^{B4} вештачка променљива која узима вредност један у периодима у којима ирегуларна компонента ($B4$) одступа од средине мање 1,5 стандардних девијација, а вредност нула у супротном случају. Њом се обезбеђује да се SI компоненте које немају екстремну I компоненту не коригују.

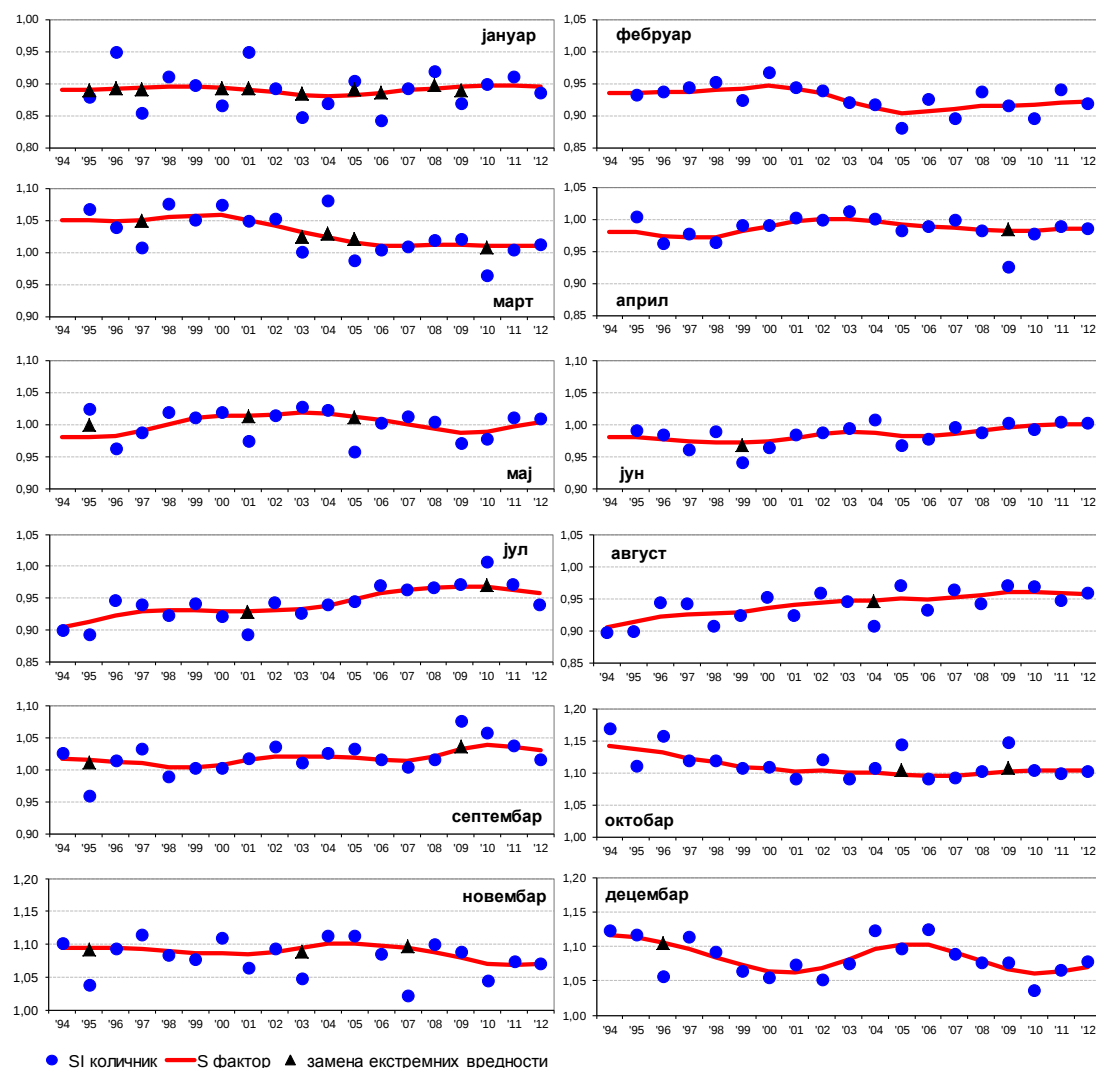
B5 Сезонски фактор

Нову, за екстреме кориговану, оцену SI користимо да оценимо S фактор. Ово радимо исто као у претходној фази, применом сезонског 3×3 покретног просека на кориговану SI серију и нормализацијом (2×12):

$$S_t^{B5} = M_{3 \times 3}^{12}(SI_t^{B4}) / M_{2 \times 12}[M_{3 \times 3}^{12}(SI_t^{B4})]$$

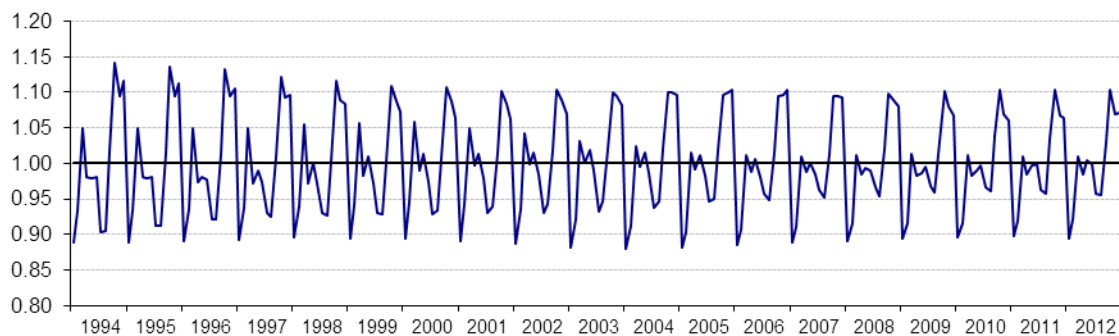
Како смо екстремне вредности кориговали, нови сезонски фактори које смо добили, овај пут нису под утицајем једнократних промена.

Слика 4.3. Оцена сезонско-ирегуларне и сезонске компоненте по месецима (B4 и B5)



Сезонски фактори промењени су у односу на прелиминарне код свих месеци код којих су замењене екстремне вредности. Када сезонске (месечне) факторе “спојимо”, добијамо серију сезонских фактора (слика 4.4) којима у следећој фази десезонизирамо серију.

Слика 4.4. Сезонска компонента (B5)

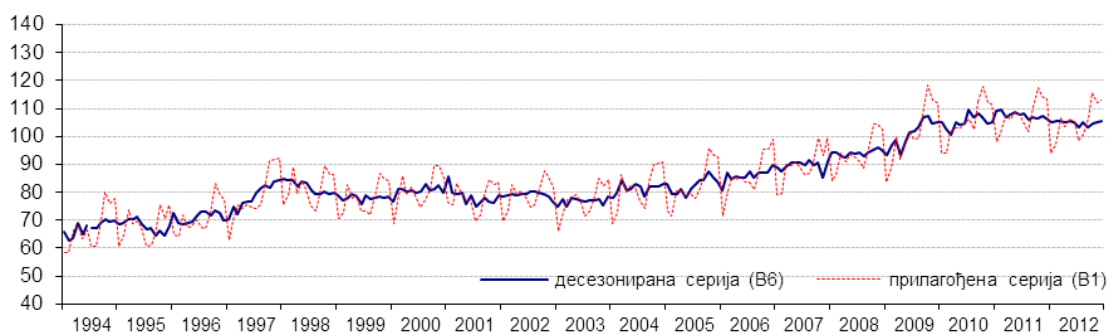


B6 Десезонирана серија

Пошто смо добили (прелиминарну) оцену сезонског фактора, дељењем прилагођене серије B1 са сезонским фактором из B5 добијамо прву прелиминарну десезонирану серију.

$$TI_t^{B6} = Y_t^{B1} / S_t^{B5}$$

У овој фази из десезониране серије још увек су искључени екстреми које смо искључили у фази пред-прилагођавања.



Екстреми које смо детектовали у B4 присутни су у десезонираној серији. Њих смо само привремено кориговали у сезонско-ирегуларној компоненти, да бисмо оценили S.

B7 Тренд – Henderson-ов филтер

Десезонирана серија коју смо добили у претходној фази у себи садржи тренд и ирегуларну компоненту. У овој фази циљ нам је да издвојимо тренд, користећи *Henderson-ов филтер*. Овај филтер је заправо пондерисани покретни просек реда 9, 13, 23, итд¹², који има ту карактеристику да када га регресирамо на параболу другог или трећег реда, регресиона оцена се савршено поклапа са самим филтером¹³. Пондере филтера израчунао је *Henderson* (1916), тако да филтерисана серија има минималну суму квадрата треће диференце.

У фази B7 тренд оцењујемо применом 9-чланог или 13-чланог *Henderson*-овог пондерисаног покретног просека на десезонирану серију. Који ћемо од ова два покретна просека применити зависи од односа ирегуларне и тренд компоненте. Већи ред филтера више изглачава серију, односно мање је под утицајем привремених

¹² Програм X-12 по аутоматизму врши избор између филтера реда 9, 13, 23, с тим што је могуће спецификати и филтере вишег реда.

¹³ National statistics (UK). "Methodology of the Monthly Index of Services Annex D: The X11ARIMA Seasonal Adjustment Program"

искока. Стога, што је већи релативан значај ирегуларне компоненте наспрам тренда, потребно је применити дужи *Henderson*-ов филтер (у овом случају реда 13).

Да бисмо оценили релативан значај две компоненте, потребна нам је њихова прелиминарна оцена. Зато, најпре оцењујемо прелиминаран тренд применом 13-чланог *Henderson*-овог филтера (H_{13}):

$$T_t^{B7p} = H_{13}(TI_t^{B6})$$

где p у суперскрипту означава да је у питању прелиминарна оцена. Пондери 13-чланог *Henderson*-овог филтера (H_{13}) приказани су у табели 4.2.

Табела 4.2. Пондери 13-чланог *Henderson*-овог филтера

доцња	0	1 i -1	2 i -2	3 i -3	4 i -4	5 i -5	6 i -6
пондер	0,24006	0,21434	0,14736	0,06549	0	-0,02786	-0,01935

Прелиминарну ирегуларну компоненту добијамо као резидуал, дељењем десезониране серије и тренда:

$$I_t^{B7p} = TI_t^{B6} / T_t^{B7p}$$

На основу ових оцена T и I добијамо просечне месечне апсолутне процентуалне промене ових компоненти:

$$\bar{T} = \frac{1}{n-1} \sum_{t=2}^n |T_t / T_{t-1} - 1|,$$

$$\bar{I} = \frac{1}{n-1} \sum_{t=2}^n |I_t / I_{t-1} - 1|,$$

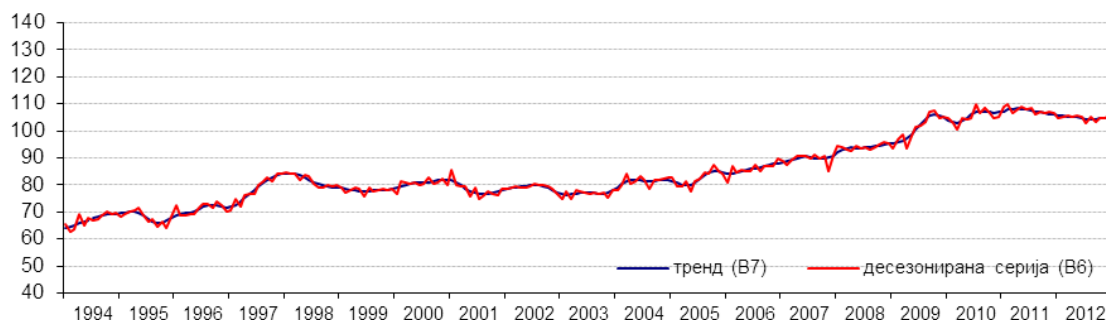
одакле добијамо однос: $I/\bar{T}$, на основу кога се одређује ред филтера:

$\bar{I}/\bar{T} < 1$ примењујемо 9-члани *Henderson*-ов филтер

$\bar{I}/\bar{T} > 1$ примењујемо 13-члани *Henderson*-ов филтер

У нашем случају, добили смо $I/\bar{T} = 3,11$, што значи да тренд оцењујемо применом *Henderson*-овог филтер реда 13 (односно задржавамо прелиминарни тренд, који је иначе био реда 13), па имамо:

$$T_t^{B7} = T_t^{B7p} = H_{13}(TI_t^{B6})$$

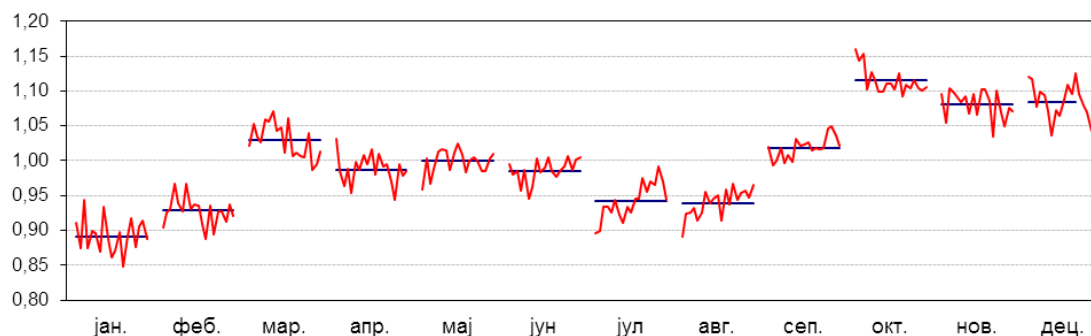


B8 Сезонско-ирегуларна компонента

Дељењем прилагођене оригиналне серије ($B1$) трендом ($B7$) добијамо нову оцену SI компоненте:

$$SI_t^{B8} = Y_t^{B1} / T_t^{B7}$$

Слика 3.5. Сезонско-ирегуларна компонента ($B8$) по месецима



B9 Замена екстремних вредности SI компоненте

У овом кораку поново коригујемо екстремне вредности SI компоненте као у $B4$. Процедура је иста, с том разликом што за оцену S компоненте (уз нормализацију), уместо 3×3 , користимо 3×5 сезонски филтер, који је еквивалентан седмочланом пондерисаном покретном просеку:

$$M_{3 \times 5}^{12}(SI_t^{B8}) = \frac{1}{15} SI_{t-36}^{B8} + \frac{2}{15} SI_{t-24}^{B8} + \frac{3}{15} SI_{t-12}^{B8} + \frac{3}{15} SI_t^{B8} + \frac{3}{15} SI_{t-12}^{B8} + \frac{2}{15} SI_{t-24}^{B8} + \frac{1}{15} SI_{t-36}^{B8},$$

У поступку идентификације екстремних вредности, чији је поступак исти као у $B4$, овај пут је као екстремно идентификовано 33 опсервације (слика 3.6). И у овој итерацији, јануар је имао највише, 9 искакања. Интересантно је да фебруар, који у $B4$

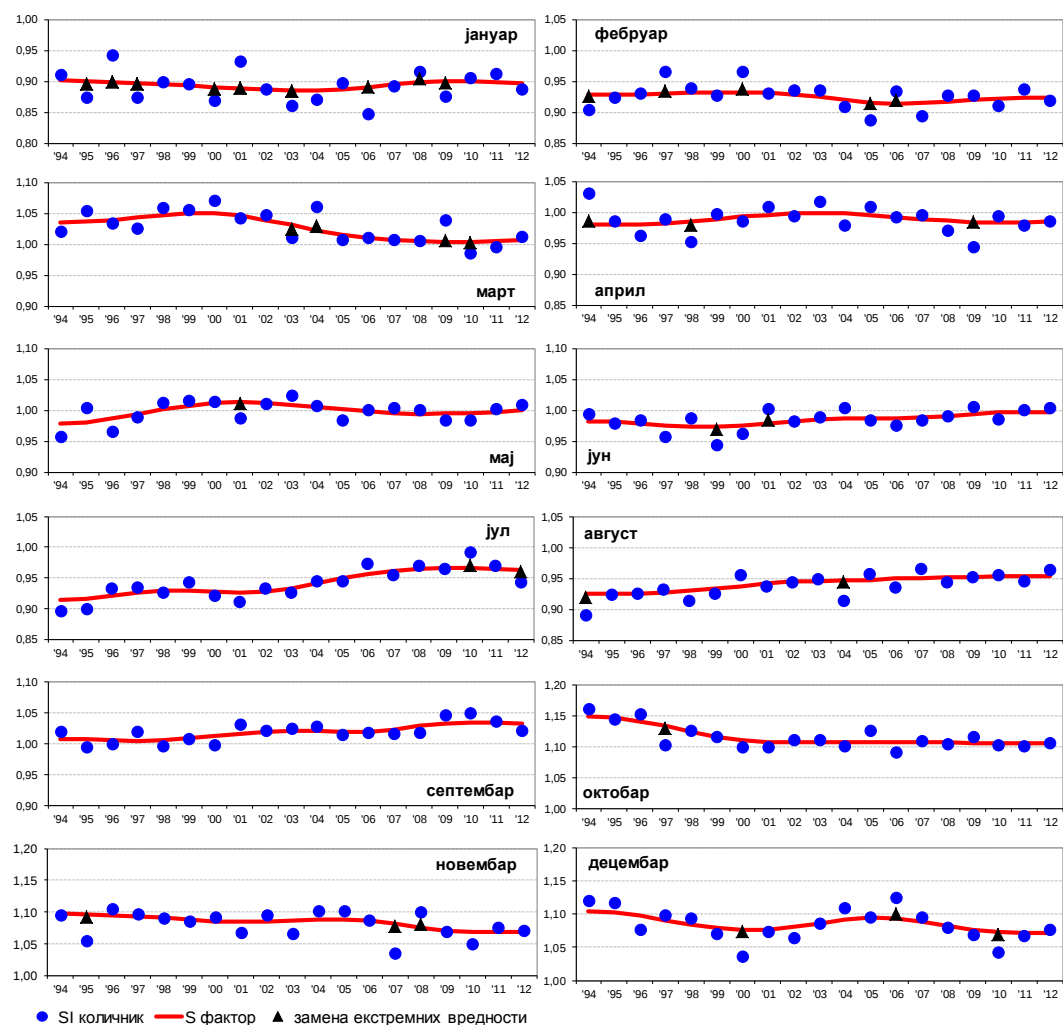
није имао ниједно искакање, сада има четири. Замењујући екстреме, добили смо нову оцену SI компоненте $SI_t^{(B9)}$.

B10 Сезонски фактор

Поступак оцене S компоненте сличан је као у $B5$ осим што уместо 3×3 примењујемо 3×5 сезонски покретни просек:

$$S_t^{B10} = M_{3 \times 5}^{12}(SI_t^{B9}) / M_{2 \times 12}[M_{3 \times 5}^{12}(SI_t^{B9})]$$

Слика 3.6. Оцена сезонско-ирегуларне ($B9$) и сезонске компоненте ($B10$) по месецима

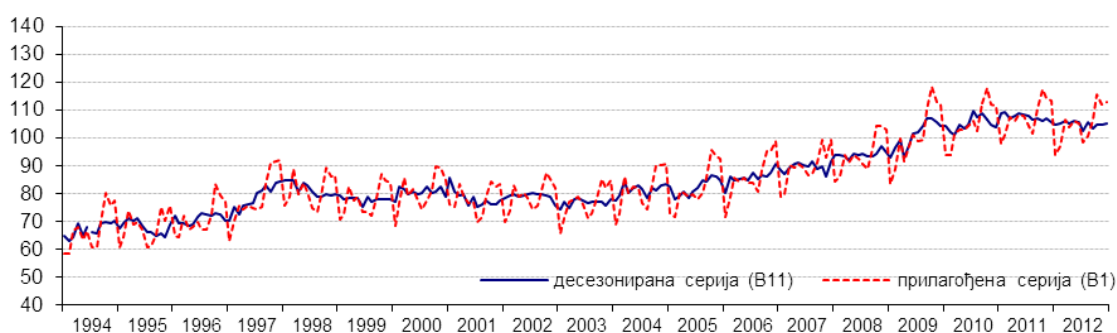


Примена већег реда филтера чини да појединачне вредности SI мање утичу на оцену S , па стога она у овом случају показује мање осцилације него што је то био случај у $B4$, када смо користили 3×3 покретни просек.

B11 Десезонирана серија

Дељењем прилагођене оригиналне серије новим сезонским фактором добијамо нову оцену десезониране серије.

$$TI_t^{B11} = Y_t^{B1} / S_t^{B10}$$



B13 Ирегуларна компонента

Дељењем десезониране серије ($B11$) трендом ($B7$) добијамо оцену ирегуларне компоненте:

$$I_t^{B13} = TI_t^{B11} / T_t^{B7}$$

B17 Прелиминарни пондери за корекцију ирегуларне компоненте

У овој фази идентификујемо екстремне вредности ирегуларне компоненте из $B13$, а затим им придружујемо одговарајуће пондере. Поступак је исти као када смо модификовали екстремне вредности SI компоненте у $B4$ и $B9$: рачунамо (петогодишње) стандардне девијације, рачунамо одступања (у стандардним девијацијама) сваке вредности I компоненте од средине и на крају придружујемо пондере свакој вредности у зависности од интензитета одступања.



На основу горе описаног метода идентификовано је 11 вредности ирегуларне компоненте које одступају више од 2,5 стандардних девијација и 21 вредност које одступају између 1,5 и 2,5 стандардних девијација. Њихове пондере рачунамо на исти начин као у B4.

B20 Корекције екстремних вредности ирегуларне компоненте

Последњи корак B фазе је оцена коефицијената за корекцију екстремних вредности на основу пондера из B17:

$$k_t^{B20} = I_t^{B13} / [1 + p_t^{B17} (I_t^{B13} - 1)].$$

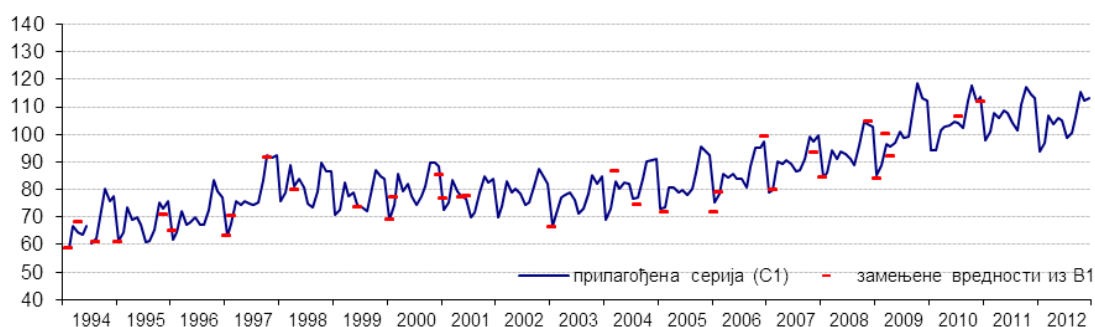
Оцена екстремних вредности представља и крајњи резултат прве, B фазе X-11 метода.

4.2. C фаза: Финална оцена екстремних вредности

C1 Прилагођена серија за екстремне вредности

Први корак у C фази је корекција прилагођене оригиналне серије из B1. Дата серија се дели коефицијентима из B20, чиме добијамо серију са коригованим екстремним вредностима.

$$Y_t^{C1} = Y_t^{B1} / k_t^{B20}$$



Наредни кораци су исти као у *B* фази. Графички прикази прелиминарних серија из *C* фазе само се незнатно разликују од одговарајућих из *B* фазе, па их зато овде изостављамо.

C2 Тренд

$$T_t^{C2} = M_{2 \times 12}(y_t^{C1})$$

C4 Сезонско-ирегуларна компонента

$$SI_t^{C3} = y_t^{C1} / T_t^{C2}$$

C5 Сезонски фактор -- 3×3 покретни просек и нормализација

$$S_t^{C5} = M_{3 \times 3}^S(SI_t^{B9}) / M_{2 \times 12} [M_{3 \times 3}^S(SI_t^{C4})]$$

C6 Десезонирана серија

$$TI_t^{C6} = y_t^{C1} / S_t^{C5}$$

C7 Тренд компонента

Поступак који примењујемо у овом кораку сличан је као у *B7*, с том разликом што се овде као једна од могућности уводи и 23-члани пондерисани покретни просек. Избор такође доносимо на основу $\bar{I}/\bar{T}$ односа:

$\bar{I}/\bar{T} < 1$ примењујемо 9-члани *Henderson*-ов филтер,

$1 < \bar{I}/\bar{T} < 3,5$ примењујемо 13-члани *Henderson*-ов филтер,

$\bar{I}/\bar{T} > 3,5$, примењујемо 23-члани *Henderson*-ов филтер.

У нашем примеру након прве прелиминарне оцене тренда и ирегуларне компоненте добили смо $\bar{I}/\bar{T} = 1,97$, што значи да смо и овој фази тренд оценили применом 13-чланог *Henderson*-овог филтера на десезонирану серију:

$$T_t^{C7} = H_{13}(TI_t^{C6})$$

C9 Сезонско-ирегуларна компонента

$$SI_t^{C9} = Y_t^{C1} / T_t^{C7}.$$

Овај пут нема корекције екстремних вредности SI количника као у $B8$ и $B9$.

C10 Сезонски фактор - 3×5 покретни просек и нормализација

$$S_t^{C10} = M_{3 \times 5}^{12}(SI_t^{C9}) / M_{2 \times 12}^{12}[M_{3 \times 5}^{12}(SI_t^{C9})]$$

C11 Десезонирана серија

$$TI_t^{C11} = Y_t^{B1} / S_t^{C10}$$

C13 Ирегуларне компоненте

$$I_t^{C13} = TI_t^{C11} / T_t^{C7}$$

C17 Прелиминарни пондери за корекцију ирегуларне компоненте

С фаза (као и B) завршава одређивањем коефицијената за корекцију екстремних вредности. У овом кораку, на $C13$ примењујемо исти алгоритам за детектовање екстремних вредности као у $B17$. Овај пут је идентификовано 32 екстремних вредности: 18 које одступају више од 2,5 стандардне девијације, и 14 које одступају 1,5-2,5 стандардних девијација.



C20 Корекције екстремних вредности ирегуларне компоненте

Као крајњи резултат овог дела, исто као у *B* фази, добијамо корективне коефицијенте за екстремне вредности:

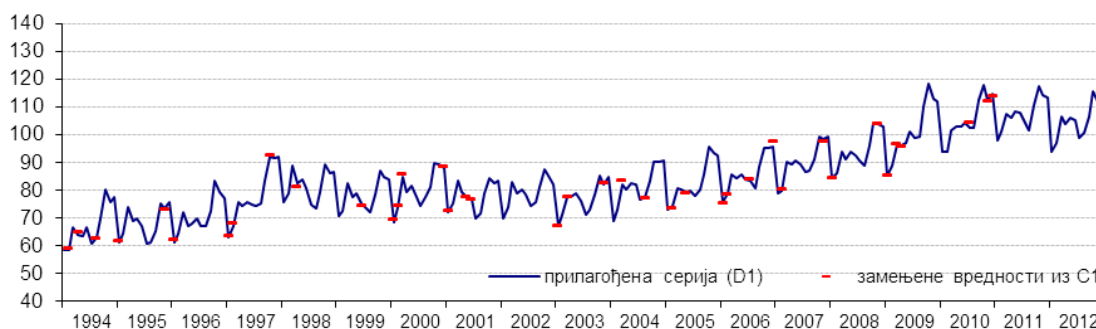
$$k_t^{C20} = I_t^{C13} / [1 + p_t^{C17}(I_t^{C13} - 1)]$$

4.3. D фаза: Финална оцена компоненти

D1 Оригинална прилагођена серија са корекцијом екстремних вредности

Слично као у *C1*, дељењем оригиналне прилагођене серије из *B1* (не *C1*) корективним коефицијентима из *C20*, добијамо прилагођену серију са корекцијом екстремних вредности (њих 24).

$$Y_t^{D1} = Y_t^{B1} / k_t^{C20}$$



Кораци *D2-D7* идентични су као *C2-C7*, с тим што се примењују на серију *D1*, па стога без појашњавања за ове кораке наводимо формуле, такође без графичког приказа.

D2 Тренд – 12-месечни центрирани покретни просек

$$T_t^{D2} = M_{2 \times 12}(Y_t^{D1})$$

D4 Сезонско-ирегуларна компонента

$$SI_t^{D4} = Y_t^{D1} / T_t^{D2}$$

D5 Сезонски фактор – 3×5 покретни просек и нормализација

$$S_t^{D5} = M_{3 \times 5}^{12}(SI_t^{D4}) / M_{2 \times 12}[M_{3 \times 5}^{12}(SI_t^{D4})]$$

D7 Тренд – Henderson-ов филтер (9, 13, 23)

Применом истог поступка као C7, и овај пут је изабран филтер реда 13 ($I/T = 1,8$).

$$T_t^{D7} = H_{13}(TI_t^{D6})$$

D8 Коначна немодификована SI компонента

$$SI_t^{D8} = Y_t^{B1} / T_t^{D7}$$

Тестови сезоналности

На основу оцене SI компоненте из D8 спроводе се три теста сезоналности: тест стабилне сезоналности, непараметарски *Kruskal-Wallis-ов тест стабилне сезоналности* и тест покретне сезоналности¹⁴. Први тест смо већ објаснили у B1.

Kruskal-Wallis-ов тест стабилне сезоналности, као и претходни, базира се на третирању оцена SI количника по сезонама као 12 независних узорака. Статистика овог теста је:

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{12} \frac{R_i^2}{n_i} - 3(N+1),$$

где је R_i , $i = 1, \dots, 12$, сума рангова сезоне i , при чему SI количнике рангирамо за цео узорак заједно (а не одвојено по сезонама).

Да бисмо разумели суштину овог теста, претпоставимо да серија има само две године опсервација, са по две сезоне. Уколико серија има изражену сезону, за очекивати је да ће једна сезона увек имати веће SI количнике од друге, па ће постојати изражена разлика између сума рангова различитих сезона (серија *a*). У супротном, ако разлика између сезона нема, суме рангова по сезонама биће (приближно) једнаке (серија *b*).

	сезона 1	сезона 2
година 1	1	3
година 2	2	4

сума R_i 3 7
серија *a*

	сезона 1	сезона 2
година 1	1	2
година 2	4	3

сума R_i 5 5
серија *b*

¹⁴ Laidray and Quenneville (1999).

Као последица, сума (и просек) квадрата сума сезонских рангова биће већа код серија са израженим сезонским осцилацијама – у нашем примеру, код серије *a* 58, а код серије *b* 50 – па ће и вредност *Kruskal-Wallis*-ове статистике бити већа.

Трећи тест је *тест покретне сезоналности*. Док у тесту стабилне сезоналности варијансу *SI* компоненте разлажемо на међу-сезонску и унутар-сезонску (резидуалну) варијансу, овде додајемо и унутар-годишњу варијансу.

$$V^2 = V_A^2 + V_B^2 + V_R^2,$$

где је

$$V^2 = \sum_{i=1}^{12} \sum_{j=1}^N (SI_{ij} - \overline{SI}_{..})^2 \text{ - укупна сума квадрата;}$$

$$V_A^2 = N \sum_{i=1}^k (\overline{SI}_{i.} - \overline{SI}_{..})^2 \text{ - међусезонска сума квадрата;}$$

$$V_B^2 = 12 \sum_{j=1}^N (\overline{SI}_{.j} - \overline{SI}_{..})^2 \text{ - међугодишња сума квадрата;}$$

$$V_R^2 = \sum_{i=1}^{12} \sum_{j=1}^N (SI_{ij} - \overline{SI}_{i.} - \overline{SI}_{.j} + \overline{SI}_{..})^2 \text{ - резидуална сума квадрата.}$$

Нулту хипотезу да нема промене у сезонском обрасцу током година тестирамо статистиком:

$$F_M = \frac{V_B^2 / (N-1)}{V_R^2 / 11(N-1)},$$

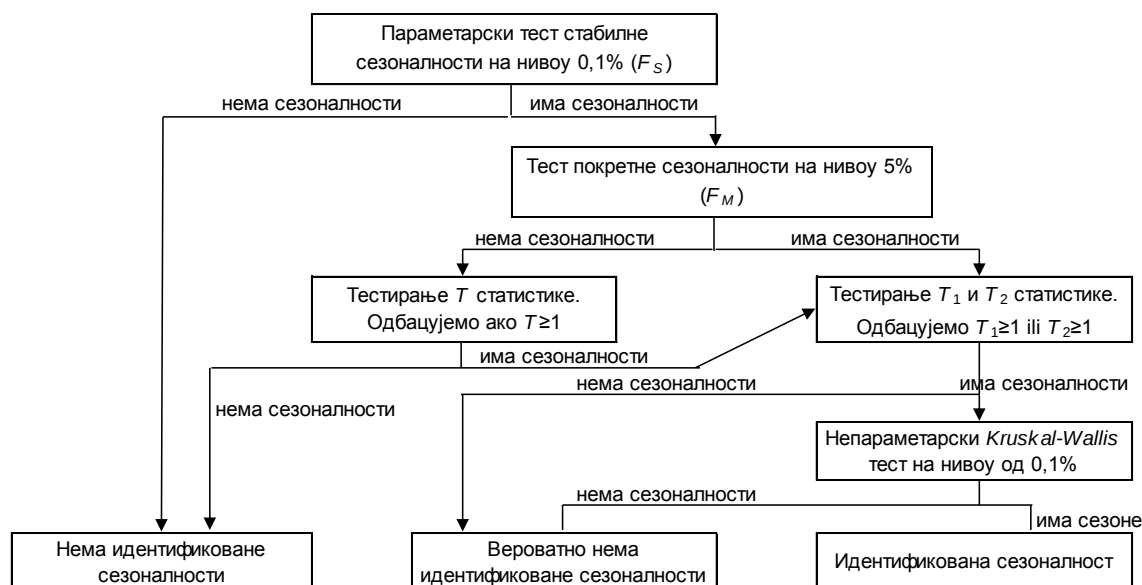
која има *F* расподелу са $(N-1)$ и $11(N-1)$ степени слободe, па уколико је израчуната вредност ове статистике изнад критичне, одбацујемо нулту хипотезу и закључујемо да је сезонски карактер серије променљив по годинама.

Ови три теста – параметарски и непараметарски тест стабилне сезоналности и тест покретне сезоналности – комбинују се у заједнички *тест идентификоване сезоналности*. Обележимо *F* статистику првог теста са F_s . Заједнички тест користи три статистике:

$$T_1 = \frac{7}{F_s}, \quad T_2 = \frac{3F_M}{F_s} \quad \text{и} \quad T = \sqrt{\frac{1}{2}(T_1 + T_2)}$$

Оцена идентификоване сезоналности серије базира се на следећем алгоритму:

Слика 4.7. Алгоритам теста идентификованог сезонског карактера



У нашем примеру добили смо следеће резултате:

Тест стабилности сезоне: Израчуната вредност F -статистике износи 183,6, што је више од критичне вредности за ризик 0,1%, па прихватамо алтернативну хипотезу о постојању стабилне сезоналности.

Kruskal-Wallis-ов тест стабилне сезоналности: Израчуната вредност $Kruskal-Wallis$ -ове статистике је 204,8, па за ризик од 1% и по овом тесту можемо да кажемо да постоји стабилна сезоналност.

Тест покретне сезоналности: Израчуната вредност F -статистике је 1,6. За ризик 5% могли бисмо да кажемо да серија има покретну сезоналност, али за ризик 1% то не би био случај.

У сваком случају, **збирни тест** (слика 4.7) показао је присуство идентификоване сезоналности.

D9 Замена екстремних вредности SI компоненте

Уместо $B1$ (као у $D8$) за оцену SI компоненте овде користитимо серију из $D1$ са модификованим екстремним вредностима:

$$SI_t^{D9} = Y_t^{D1} / T_t^{D7}$$

Ове две оцене SI компоненте (у $D8$ и $D9$) разликоваће се у месецима у којима смо извршили корекције екстремних вредности у корацима B и C . Вредности SI компоненте из $D9$ које се разликују у односу на $D8$ представљају замене екстремних вредности.

D9a Количници покретне сезоналности

На основу SI компоненте из $D9$ можемо оценити релативан значај ирегуларне и сезонске компоненте, на основу чега рачунамо количнике покретног сезонског обрасца. Поступак је сличан као у $B7$ и $C7$ када смо поредили ирегуларну и тренд компоненту. Сезонски фактор оцењујемо обичним седмочланим покретним просеком од SI_t^{D9} . Да бисмо користили симетричан покретни просек потребне су нам по три вредности са обе стране. Недостајуће податке на крајевима оцењујемо просечним вредностима прве односно последње три опсервације. Ако, на пример, имамо опсервације од x_1 до x_N , оцењујемо:

$$x_{-2} = x_{-1} = x_0 = (x_1 + x_2 + x_3)/3 \quad \text{и} \quad x_{N+1} = x_{N+2} = x_{N+3} = (x_N + x_{N-2} + x_{N-3})/3$$

Сада можемо да применимо седмочлани покретни просек да бисмо добили оцену сезонског фактора S . Ирегуларну компоненту добијамо као резидуал, дељењем SI са S . Након тога, рачунамо просечан апсолутни међугодишњи раст¹⁵ за сваки месец посебно:

$$\bar{S}_i = \frac{1}{N_i - 1} \sum_{t=2}^{N_i} |S_{i,t}/S_{i,t-1} - 1| * CS_i, \quad \bar{I}_i = \frac{1}{N_i - 1} \sum_{t=2}^{N_i} |I_{i,t}/I_{i,t-1} - 1| * FIS_i,$$

где је N_i број месеци i , а CS_i и FIS_i корективни коефицијенти који су функција броја година¹⁶:

$$CS_i = \frac{\sqrt{3}N_i}{6\sqrt{2} + (N_i - 6)\sqrt{3}}, \quad FIS_i = \frac{5\sqrt{6}N_i}{6\sqrt{149} + 5\sqrt{6}(N_i - 6)\sqrt{3}}.$$

Количник покретног сезонског обрасца за месец i рачунамо као: $RSM_i = I_i/S_i$, док је општи (заједнички) количник покретног сезонског обрасца:

$$RSM = \frac{\sum_i N_i \bar{I}_i}{\sum_i N_i \bar{S}_i}.$$

¹⁵ Месец текуће према истом месецу претходне године.

¹⁶ Приказане формуле односе се на случај када имамо 7 или више година опсервација. За краће узорке формула је другачија (видети *Laidray and Quenneville*, 1999).

D10 Финални сезонски фактор

На основу количника RSM из $D9$ бирамо одговарајући покретни просек који примењујемо на серију SI_t^{D9} . Количници из $D9$ мере релативан значај шума. Што је количник RSM већи, серија је хаотичнија, па треба изабрати већи ред покретног просека.

Ред покретног просека бира се на основу следећих критеријума:

$RSM < 2,5$ бирамо 3×3 покретни просек;

$3,5 < RSM < 5,5$ бирамо 3×5 покретни просек;

$RSM > 6,5$ бирамо 3×9 покретни просек;

Ако RSM упадне у једну од међузона за које нема решења (2,5-3,5, 5,5-6,5), коефицијент поново рачунамо са узорком скраћеним за једну годину на крајевима. Ако поново нема решења, одузимамо две године и поново рачунамо, и тако све док не дођемо до решења. Уколико ни након скраћивања узорка за пет година нема решења, бирамо 3×5 покретни просек.

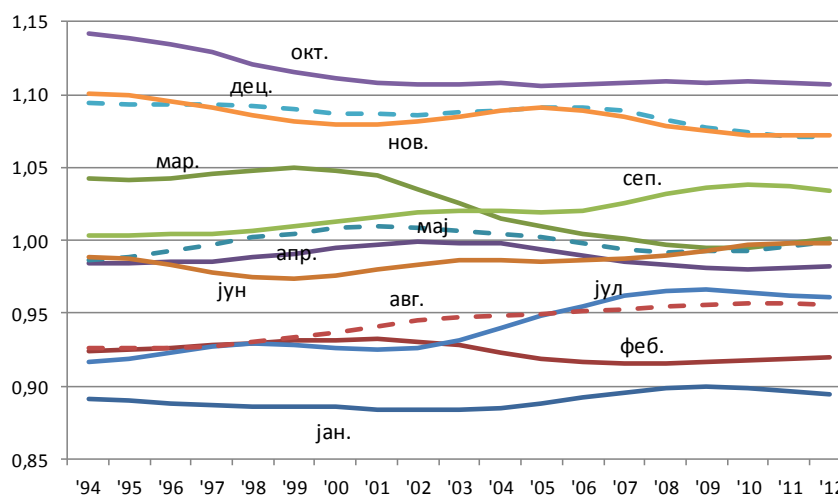
Изабрани покретни просек, затим, примењујемо на серију SI_t^{D9} , да бисмо добили S компоненту. Овај сезонски фактор нормализујемо 12-месечним покретним просеком. Сезонски фактор за будуће године прогнозира се једноставном линеарном пројекцијом на бази последње две године:

$$S_{N+1} = S_N + (S_N - S_{N-1})/2$$

У нашем примеру добили смо $RSM=5,2$, што значи да финални сезонски фактор (слика 4.8) добијамо применом 3×5 филтера на сезонско-ирегуларну компоненту из претходног корака:

$$S_t^{D10} = M_{3 \times 5}^{12}(SI_t^{D9})$$

Слика 4.8. Финални сезонски фактор (D10)



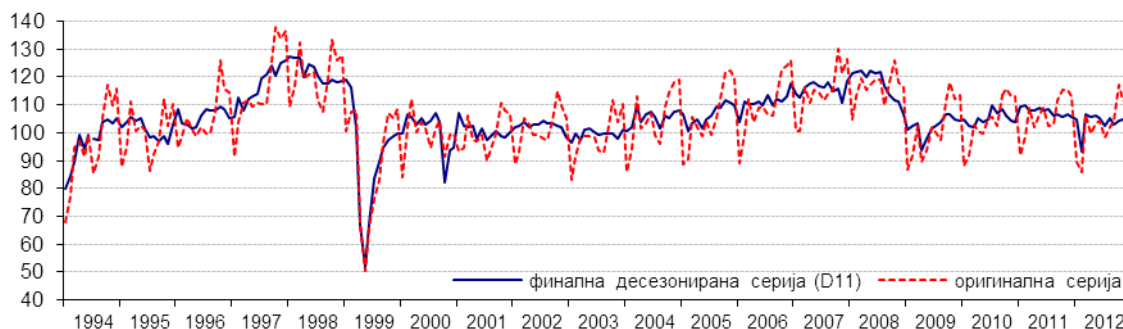
Као што можемо видети, у односу на почетак серије дошло је до промене сезоналности у правцу мањих разлика између летњих и јесењо-зимских месеци. Ово се може објаснити променом структуре годишњих одмора, тј. њиховим све већим коришћењем у јесењим и зимским месецима, на штрб летњих, мада ову претпоставку не можемо да поткрепимо конкретним подацима. Могуће је и да повећана употреба клима уређаја доприноси већој производњи електричне енергије у летњим месецима.

D11 Финална десезонирана (и декалендарисана) серија

Делећи оригиналну серију (пре пред-прилагођавања) само финалним сезонским фактором ($D10$) добили бисмо финалну десезонирану серију која би у себи укључивала календарске ефекте (радни дани и празници). У овој фази имамо опцију, коју смо ми изабрали, да из оригиналне серије, поред сезонских искључимо и календарске ефекте. Календарску компоненту K_t смо оценили у фази пред прилагођавања, па финалну десезонирану и *декалендарисану* (без календарских ефеката) серију добијамо као:

$$TI_t^{D11} = Y_t / S_t^{D10} / K_t$$

Приметимо да су у финалној десезонираној серији остали ефекти скока које смо искључили у пред-прилагођавању.



У овој фази тестирамо да ли се у ирегуларној компоненти десезониране серије задржала сезона (пожељно је да није). Ирегуларну компоненту у овој фази оцењујемо као трећу диференцу десезониране серије:

$$I_t^{D11} = \Delta_3 TI_t^{D11} = TI_t^{D11} - TI_{t-3}^{D11}$$

Затим спроводимо тестове стабилног сезонског карактера на диференцираној, тј. ирегуларној серији на исти начин као у ВЗ. Један тест спроводимо на целој серији, а један на последње три године. У нашем примеру, за целу серију добили смо $F = 0,54$, а за последње три године $F = 0,52$, и у оба случаја, на нивоу значајности од 1%, закључујемо да у резидуалима нема сезоне.

D12 Финални тренд – Henderson-ов филтер

Финални тренд добијамо применом *Henderson*-овог филтера на десезонирану серију $D1$, коју добијамо делећи прилагођену серију из $D1$ финалним сезонским фактором $S10$ (Y_t^{D1} / S_t^{D10}). Поступак и критеријуми избора реда *Henderson*-овог филтера (9, 13, 23) исти су као и у корацима $C7$ и $D7$. И у овом кораку изабран је филтер реда 13 ($\bar{I}/\bar{T}=2,1$). Уколико би се овде завршила оцена финалног тренда, он у себи не би садржао структурни лом.

Већ смо рекли да финална десезонирана серија садржи искоке које смо елиминисали у фази пред-прилагођавања. Како се десезонирана серија састоји из тренда и ирегуларне компоненте, остаје питање како расподелити искоке на ове две компоненте. Ирегуларна компонента по својој дефиницији представља одступања од тренда која могу бити само привремена. Са друге стране, тренд представља дугорочну тенденцију серије, па стога по својој природи не би требало да садржи привремене искоке. Због тога, метод *X-12-ARIMA* трајне ломове, какви су LS и RP , укључује у тренд, а привремене искоке, AO и TC , у ирегуларну компоненту.

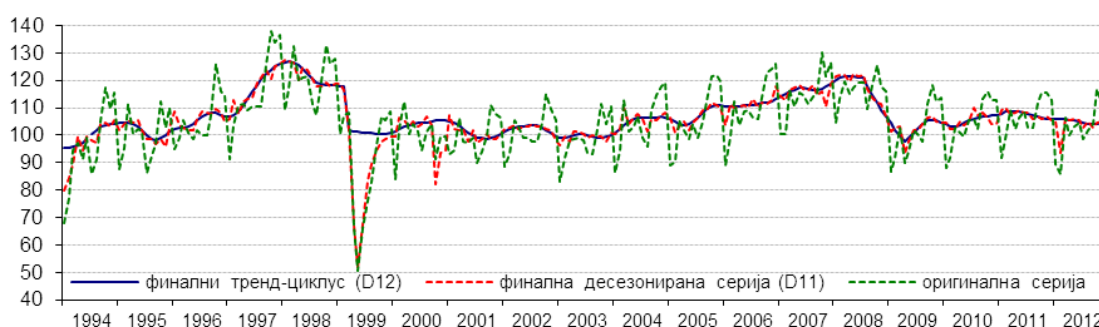
Стога је у тренд неопходно укључити ломове LS^{mar99} и $RP^{avg08-maj09}$, чије смо ефекте оценили у фази пред-прилагођавања. Овде је, слично као и код календарских ефеката, ове променљиве неопходно прилагодити мултипликативном методу, нпр:

$$efLS^{mar99} = e^{\beta_{LS} \cdot LS^{mar99}} = e^{-0,1415 LS^{mar99}}$$

$$efRP^{avg08-maj09} = e^{\beta_{RP} \cdot RP^{avg08-maj09}} = e^{-0,0286 RP^{avg08-maj09}}$$

па је финална оцена тренд-циклус компоненте

$$T_t^{D12} = H_{13}(Y_t^{D1} / S_t^{D10}) \cdot efLS^{mar99} \cdot efRP^{avg08-maj09}$$

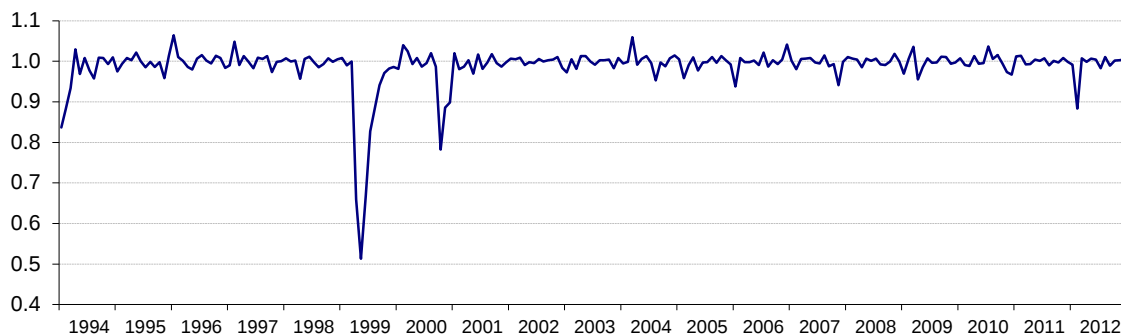


D13 Финална ирегуларна компонента

Последњи корак је оцена финалне ирегуларне компоненте, коју добијамо делећи финалну десезонирану серију финалним трендом.

$$I_t^{D13} = \frac{TI_t^{D11}}{T_t^{D12}}$$

Слика 4.9. Финална оцена ирегуларне компоненте (D13)



Како финална десезонирана серија садржи све искоке, а финални тренд само трајне (*LS* и *RP* искок), количник ове две серије, тј. финална ирегуларна компонента садржи привремене, *AO* и *TC* искоке за 1994, 1999, 2000. и 2012. годину. Календарске ефекте, K_t , искључили смо из десезониране серије, то се она не налази ни у ирегуларној компоненти.

* * *

Након што смо оценили све основне компоненте серије, оригиналну серију индустријске производње можемо да представимо као њихов производ

$$Y_t = T_t^{D12} \times S_t^{D10} \times K_t \times I_t^{D13}.$$

Како смо применили мултипликативни метод, компоненте S , K и I узимају вредности око један, и заправо дефинишу сезонска, календарска и ирегуларна (несистематска) одступања од тренда серије.

Овде је *X-11* фаза, као и цео поступак десезонирања, завршен. На крају, неопходно је проверити да ли извршено десезонирање адекватно извршено.

5. Провера адекватности десезонирања

Након што смо добили десезонирану серију, неопходно је проверити квалитет извршеног десезонирања. Постоје бројни методи провере адекватности, од којих ћемо у овом раду обрадити три: M статистике, анализу спектра и ревизије.

5.1. Дијагностички тестови – M статистике

Најчешће коришћени метод адекватности десезонирања су M статистике, њих једанаест, и збирна, Q статистика свих M статистика. Све статистике су стандардизоване тако да се сматрају задовољавајућим уколико узимају вредност мању од један и обратно.

$M1$ – Допринос ирегуларне компоненте укупној варијанси серије. Ако је овај допринос већи, теже је идентификовати и издвојити сезонску компоненту, па су закључци у погледу ове компоненте мање поуздани. Допринос ирегуларне компоненте до 10% сматра се задовољавајућим.

$M2$ – Допринос ирегуларне компоненте варијанси стационираној компоненти оригиналне серије. Као код претходне статистике, прихватљив је допринос до 10%.

$M3$ – Допринос ирегуларне компоненте променама (прелиминарне) десезониране серије. Пожељно је да овај допринос не буде превисок, јер би у супротном било тешко раздвојити тренд-циклус и ирегуларну компоненту.

$M4$ – Присуство аутокорељације ирегуларне компоненте. Валидност F -теста из $X-11$ базира се на случајној природи ирегуларне компоненте, па је пожељно да ирегуларна компонента није аутокорељисана.

$M5$ – Број месеци неопходан да би варијације у тренду надмашиле варијације у ирегуларној компоненти. Слично као $M3$, ова вредност може се користити за поређење значаја тренд и ирегуларне компоненте.

$M6$ – Годишње промене ирегуларне компоненте исувише мале у односу на годишње промене сезонске компоненте (низак I/S количник). Када је овај однос исувише мали, 3×5 покретни просек (којим филтришемо S из SI компоненте) не прати довољно флексибилно сезонска кретања.

$M7$ – Релативан допринос стабилне и покретне сезоне. Сезонски филтери у $X11$ дају најбоље резултате уколико је сезона стабилна.

$M8$ – Варијације сезонске компоненте на бази (квази)-случајних краткорочних варијација. Рачуна се као просек апсолутних годишњих промена нормализоване сезонске компоненте.

$M9$ – Варијације сезонске компоненте на бази на бази дугорочних промена. Рачуна се као просек годишњих промена нормализоване сезонске компоненте.

$M10$ – Исто што и $M8$, само за последње три године.

$M11$ – Исто што и $M9$, само за последње три године.

На крају, програм $X-12$ тестира свих 11 статистика заједно. Збирна Q статистика представља пондерисани просек статистика $M1-M11$:

$$Q = \frac{10M1 + 11M2 + 10M3 + 8M4 + 11M5 + 10M6 + 18M7 + 7M8 + 7M9 + 4M10 + 4M11}{100}$$

У нашем примеру, свих једанаест M статистика и заједничка Q статистика, узеле су вредност знатно мању од један:

$$\begin{aligned} M1 &= 0,163; M2 = 0,561; M3 = 0,510; M4 = 0,102; M5 = 0,606; M6 = 0,429; \\ M7 &= 0,181; M8 = 0,355; M9 = 0,179; M10 = 0,349; M11 = 0,300; \quad Q = 0,32 \end{aligned}$$

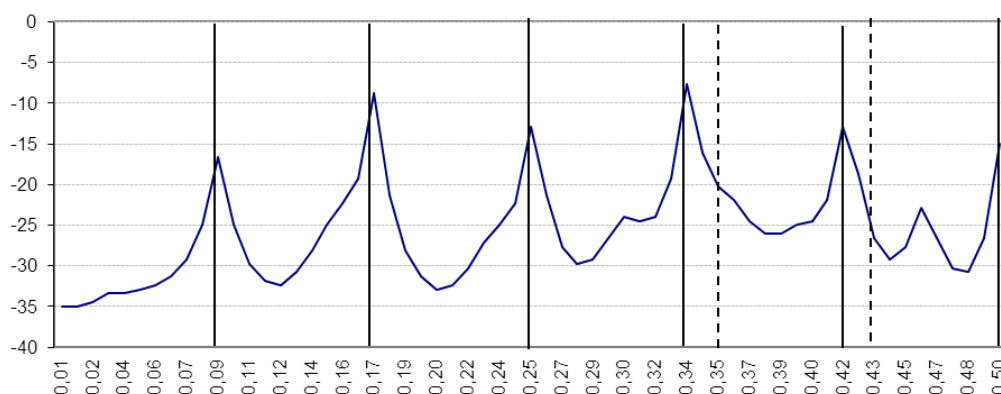
На основу свих критеријума, према томе, можемо да закључимо да је десезонирање адекватно извршено.

5.2. Анализа спектралних функција

Други метод провере адекватности десезонирања, о коме ћемо говорити у овом раду, је анализа спектралних функција. За очекивати је да серија пре десезонирања има спектралне врхове на сезонским фреквенцијама. Уколико је десезонирање адекватно извршено, десезонирана (TI) серија и ирегуларна компонента десезониране серије не би требало да имају спектралне врхове на сезонским фреквенцијама. Како искоци могу значајно утицати на спектралну функцију, посматрамо серије кориговане за искоке.

У нашем примеру упадљив је значај сезонских фреквенција на спектралној функцији оригиналне серије (диференциране и кориговане за искоке) пре десезонирања (слика 5.1).

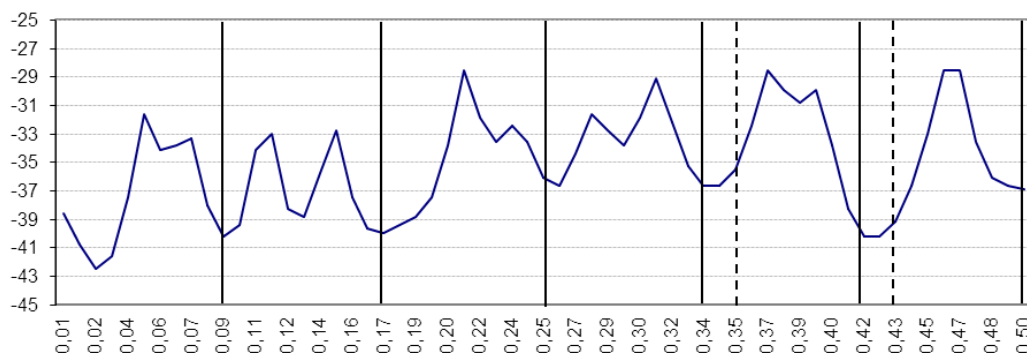
Слика 5.1. Функ. спектр. густине диференце оригиналне серије кориговане за искоке



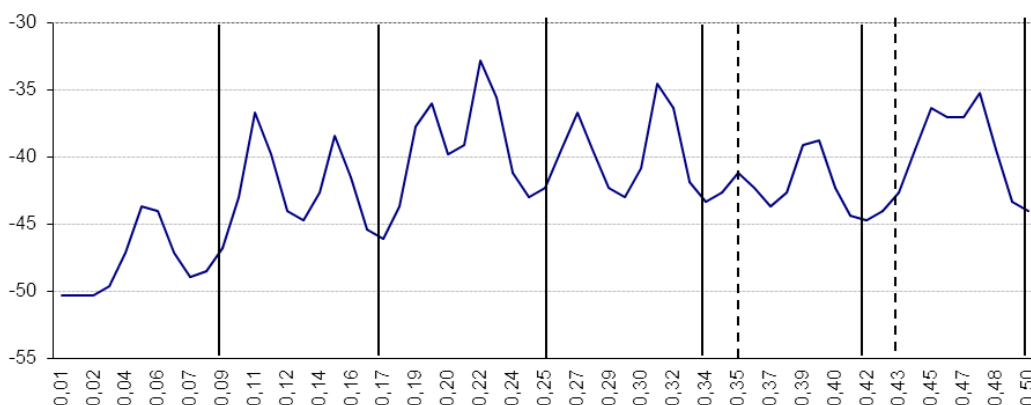
Међутим, спектар диференциране серије и њене ирегуларне компоненте немају ту карактеристику (слике 5.2. и 5.3). Можемо да закључимо да је сезонска компонента

одстрањена из серије, па је и на основу овог критеријума десезонирање адекватно извршено.

Слика 5.2. Функ. спектр. густине диференце десезониране серије кориговане за искоке



Слика 5.3. Функција спектралне густине ирегуларне компоненте кориговане за искоке



5.3. Ревизије

Као што смо видели у претходном делу, десезонирање методом *X-11* у највећој мери базира се на употреби симетричних филтера. Примена ових филтера на крајње вредности у серији захтева неку оцену будућих вредности, за шта користимо *regARIMA* модел, или употребу асиметричних филтера тамо где недостају будуће (оцењене) вредности. Када пристигне нови податак, десезонирање се врши поново, с тим што овај пут, када оцењујемо десезонирану серију (или тренд) за период који је раније био крајњи, имамо једну стварну будућу вредност која се по правилу разликује од раније оцене. Ово утиче на ревизију десезониране серије и тренда крајњих периода. Ревизија се наставља кад пристигне и следећи податак и тако редом. Што су ове ревизије мање, можемо да претпоставимо да је оцена на крају серије прецизнија.

Ако се $TI_{t|n}$ обележимо десезонирану серију за период t ($t = 1, \dots, T$) на основу серије $y_1, y_2, \dots, y_t, \dots, y_n$, ($t \leq n \leq T$) текућа десезонирана оцена серије за период t била би $TI_{t|t}$, а финална оцена $TI_{t|T}$. Финална ревизија текуће оцене би била процентуална промена финалне десезониране оцене ($t|T$) односу текућу оцену ($t|t$)

$$R_{t|T} = (TI_{t|T} - TI_{t|t}) / TI_{t|t} ,$$

а ревизија оцене након d периода

$$R_{t|t+d} = (TI_{t|t+d} - TI_{t|t}) / TI_{t|t}$$

У табели 5.1. приказане просечне апсолутне ревизије десезониране серија по годинама (од 2006.) и за цео период. Након једног периода оцене се у просеку ревидирају за 0,23%. Као што је за очекивати, ревизија се повећава са бројем периода, па је након две године просечна ревизија 0,77%. Текућа оцена у просеку одступа од финалне за 0,73%. Посматрано по годинама, ревизије у кратком року су најмање за последњу годину серије, 2012, а највеће за 2009. Ревизије након 12 и 24 периода највеће су за 2008, вероватно због кризе која је уследила крајем те године.

Табела 5.1. Ревизија десезониране серије, по годинама
(просечна апсолутна процентуална одступања)

	Финална ревизија	након 1 периода	након 2 периода	након 3 периода	након 12 периода	након 24 периода
2006.	0,72	0,19	0,26	0,24	0,43	0,84
2007.	0,94	0,25	0,30	0,32	0,73	0,76
2008.	1,12	0,28	0,51	0,66	1,13	1,13
2009.	0,78	0,36	0,64	0,76	0,50	0,80
2010.	0,47	0,26	0,34	0,36	0,43	0,46
2011.	0,35	0,17	0,25	0,28	0,23	*****
2012.	0,16	0,08	0,12	0,17	*****	*****
Укупно	0,73	0,23	0,33	0,38	0,59	0,77

Ако ревизије посматрамо по месецима (табела 5.2), приметићемо да су оне највеће за јануар, нарочито за дуже временске интервале (12 и 24 месеца).

Табела 5.2. Ревизија десезониране серије, по месецима
(просечна апсолутна процентуална одступања)

	Финална ревизија	након 1 периода	након 2 периода	након 3 периода	након 12 периода	након 24 периода
Јан	1,10	0,30	0,46	0,48	0,96	1,33
Феб	0,67	0,28	0,42	0,47	0,59	0,71
Мар	1,42	0,22	0,35	0,29	0,77	1,12
Апр	0,47	0,16	0,15	0,29	0,34	0,44
Мај	0,70	0,12	0,33	0,30	0,52	0,75
Јун	0,53	0,29	0,28	0,30	0,35	0,49
Јул	0,79	0,21	0,27	0,36	0,50	0,77
Авг	0,76	0,19	0,32	0,38	0,54	0,60
Сеп	0,69	0,20	0,31	0,42	0,53	0,81
Окт	0,42	0,17	0,32	0,38	0,62	0,58
Нов	0,50	0,29	0,41	0,46	0,66	0,69
Дец	0,72	0,28	0,36	0,48	0,70	0,93
Укупно	0,73	0,23	0,33	0,38	0,59	0,77

Исту врсту анализе можемо да урадимо и за оцене тренда серије (табела 5.3). Упадљиво је да су ревизије тренда знатно израженије него ревизије десезониране серије, нарочито за мање доцње, што генерално и јесте разлог због кога се у анализи текуће привредне активности користе десезониране вредности, а не тренд, иако би се за прво могла ставити замерка да због ирегуларне компоненте коју садржи није адекватна мера.

Табела 5.3. Ревизије тренда серије, по годинама
(просечна апсолутна процентуална одступања)

	Финална ревизија	након 1 периода	након 2 периода	након 3 периода	након 12 периода	након 24 периода
2006.	0,88	0,76	0,89	0,85	0,77	0,86
2007.	1,23	0,87	1,07	1,05	1,09	1,06
2008.	1,84	1,04	1,52	1,77	1,66	1,92
2009.	1,81	1,4	2,01	2,17	1,94	1,76
2010.	1,15	1,01	1,18	1,24	1,1	1,02
2011.	0,52	0,43	0,53	0,58	0,52	*****
2012.	0,38	0,34	0,35	0,44	*****	*****
Укупно	1,09	0,81	1,00	1,08	1,11	1,18

Иако је боље да су ревизије што мање, не постоји неки јединствен тест или критеријум којим се утврђује да ли је степен ревизије задовољавајући или не. Ревизије једноставно треба анализирати када дефинишемо претпоставке десезонирања и када поредимо десезонирање на бази различитих претпоставки.

Уколико десезонирање не би задовољило горње критеријуме, било би неопходно поново га извршити уз измењене претпоставке. Напомињемо да је приказани пример процеса десезонирања представља само последњу итерацију. Десезонирање смо урадили неколико пута, док нисмо добили задовољавајуће критеријуме адекватности.

6. Закључак

У овом раду смо детаљно, корак по корак, анализирали десезонирање серије индустријске производње Србије методом *X-12*. Најпре смо у фази пред-прилагођавања серију прочистили од највећих искока и календарских ефеката и прогнозирали је две године унапред. Као највеће искоке *regARIMA* је аутоматски оценио периоде: крај хиперинфлације 1994. (на почетку серије), НАТО бомбардовање 1999, петооктобарске политичке промене 2000 и талас захлађења у фебруару 2012. Поред тога, ми смо специфицирали лом рампу за почетак светске економске кризе (период јул 2008. – мај 2009.), који се показао статистички значајним.

RegARIMA је погодан за статистичку оцену појединих догађаја на кретање серије. НАТО бомбардовање је, поред снажног негативног утицаја у току четворомесечног рата, имало и трајан негативан утицај на ниво серије 13%. Оцењени ефекат светске економске кризе је такође трајан и креће се око 23%. У преосталим случајевима, ефекат је привремен (*AO* или *TC*), а, примера ради, талас захлађења у фебруару 2012. (*AO*) је довео до једнократног пада индустријске производње од 12%. Статистички значајним се показао и ефекат броја радних дана и празника, а за оба регресора оцењена (прерачуната) еластичност се креће око 1,5. Другим речима, један радни дан мање у месецу, због календара или празника, за последицу има 1,5% мању производњу у том месецу.

Серија прочишћена од искока и календарских ефеката ушла је у наредну фазу – *X-11*. У раду смо приказали све кораке ове фазе, која се у највећој мери састоји из низа пондерисаних покретних просека којима се наизменично оцењују компоненте серије: тренд, сезонска и ирегуларна. Поред тога, иако је у фази пред-прилагођавања серија прочишћена од највећих искока, и у фази *X-11* се у више корака (*B4*, *B17*, *B20*, *C4*, *C17*, *C20*) тестира се присуство екстремних вредности у *S* и *SI*, које се затим замењују другим (не-екстремним) вредностима, да не би утицале на оцену основних компоненти серије.

Као месец са највећим бројем искакања у свим овим фазама показао се јануар. То је вероватно повезано с варијабилним утицајем који јануарски празници имају на индустријску производњу, а у одређеној мери и са временским условима, с обзиром да таласи снажног захлађења могу проузроковати привремене прекиде производње у појединим областима. Мада смо у фази пред-прилагођавања обухватили утицај празника и серију прочистили од истих, ефекат повезивања празника није лако обухватити регресијом и он се вероватно у одређеној мери задржао у пречишћеној серији која је ушла у *X-11* фазу.

У току фазе *X-11* се у пар наврата тестира постојање сезоналности и њеног карактера. Тестови базирани на претпоставци стабилне сезоналности (анализа варијансе и *Kruskal-Wallis*-ов тест) недвосмислено потврђују постојање сезоналности за ниво значајности од 1%. Интересантно је да тест анализе варијансе базиран на претпоставци покретне сезоналности такође показује присуство сезоналности, али нешто мање значајно него у претходним случајевима, на нивоу значајности 5%.

Промена сезоналности, у правцу мањих разлика између летњих и јесењо-зимских месеци, у посматраном периоду приметна је и на графичком приказу оцењених финалних месечних сезонских фактора. Јесењо-зимски месеци (октобар-март) током времена имају све ниже сезонске факторе, а летњи (јул-септембар) све више (мање негативне). Ово се може објаснити променом структуре годишњих одмора, тј. њиховим све већим коришћењем у јесењим и зимским месецима, науштрб летњих, мада ову претпоставку не можемо да поткрепимо конкретним подацима. Могуће је и да повећана употреба клима уређаја доприноси већој производњи електричне енергије у летњим месецима.

У последњим корацима X-11 фазе, у финалне компоненте серије враћају се искоци оцењени у фази пред-прилагођавања. Тако су трајни искоци повезани с НАТО бомбардовањем и светском економском кризом (LS^{mar99} и $RP^{avg08-maj09}$) укључени у финални тренд, док финална ирегуларна компонента садржи привремене искоке (TC^{jan94} , AO^{apr99} , TC^{okt00} и AO^{feb12}). У овом раду, финална десезонирана серија очишћена је како од сезонских, тако и од календарских ефеката оцењених у RegARIMA фази, чиме се добија реалнија слика о кретању индустријске производње.

У последњем делу смо проверили адекватност примењеног десезонирања индустријске производње. M-статистике, њих 11, којима се тестирају карактеристике ирегуларне и сезонске компоненте, су потврдиле адекватност десезонирања. Такође, спектрална анализа није показала је присуство сезонске и календарске фреквенције у десезонираној серији, што показује да су из серије успешно одстрањене ове две компоненте.

На крају смо тестирали у коликој мери се оцењена десезонирана серија ревидира додавањем новог податка и поновним десезонирањем. Као што је за претпоставити, ревизије расту са порастом временског помака. Тако се додавањем једне опсервације, последња десезонирана вредност (пре додавања) у просеку (апсолутно) промени за 0,23%, након три месеца 0,38%, а након две године 0,77%. Интересантно је да се и овде јануар појављује као месец који је најмање „поуздан“, тј. као месец за кога су ревизије десезониране серије највеће, нарочито након годину дана и више. То је такође вероватно повезано с великим бројем екстрема серије у овом месецу.

Као што је за очекивати, ревизије тренда су знатно веће него ревизије десезониране серије, нарочито за мање доцње (1,08% након три месеца). То генерално и јесте разлог због кога се у анализи текуће економске активности користе десезониране вредности, а не тренд, иако би се првом могла ставити замерка да због ирегуларне компоненте коју садржи није адекватан показатељ.

Библиографија

- Cleveland W.S. and S.J. Devlin. 1982. "Calendar Effects in Monthly Time Series: Modeling and Adjustment". *Journal of the American Statistical Association*, Vol. 77, No. 379
- DAGUM E.B. 1988. "The X11ARIMA/88 Seasonal Adjustment Method, Foundations and User's Manual". *Statistics Canada technical report*
- Deutsche Bundesbank. 1999. "The changeover from seasonal adjustment method Census X-11 to Census X-12-ARIMA". *Deutsche Bundesbank Monthly Report*, September 1999
- Dosse J. and Plannas C. 1996. "Pre-adjustment in Seasonal Adjustment Methods: A Comparaison of REGRAIMA and TRAMO". *Eurostat working group document n° D3/SA/07*.
- European Central Bank. 2000. "ECB Seasonal adjustment of monetary aggregates and HICP for the euro area". <http://www.ecb.int/pub/pdf/other/sama0008en.pdf>
- Hood C.C. and Monsell B.C. 2002. "Getting Started with X-12-ARIMA Input Files on Your PC – DRAFT". <http://www.census.gov/ts/papers/gettingstartedx12.pdf>
- Laidray, D., and B. Quenneville. 1999. "Understanding the X11 Method: The Various Tables". *Working Paper, Statistics Canada*.
- Lothian J., Morry M. 1978. "A Set of Quality Control Statistics for the X11-ARIMA88 Seasonal Adjustment Method". *Working paper 78-10-005, Statistics Canada*
- National statistics (UK). "Methodology of the Monthly Index of Services Annex D: The X11ARIMA Seasonal Adjustment Program".
http://www.statistics.gov.uk/iosmethodology/downloads/Annex_D_The_X11ARIMA_seasonal_adjustment_method.pdf
- Shiskin J., Young A., Musgrave J. 1967. "The X11 Variant of the Census Method II Seasonal Adjustment Program". *Technical Paper n°15, US Bureau of the Census, Washington*.
- U.S. Census Bureau. 2011, "X-12-ARIMA Reference Manual".
<http://www.census.gov/ts/x12a/final/temp/x12adocV03.pdf>