



Statistiska centralbyrån Statistics Sweden

Vägledning för säsongrensning



Innehåll

1	Syftet med denna vägledning	3
1.1	Två relaterade dokument	3
1.2	Från teoretiskt ramverk till praktisk tillämpning	3
1.3	Vägledningens upplägg	4
2	Principer och metoder	4
2.1	Förberedelser	5
2.1.1	Allmänt.....	5
2.1.2	Konsistensfrågor	5
2.2	Outlierhantering.....	8
2.2.1	Allmänt.....	8
2.2.2	Exempel med outlierhantering	10
2.3	Kalenderkorrigering	11
2.4	Val av ARIMA-modell.....	13
2.4.1	Allmänna hållpunkter	13
2.4.2	Tillvägagångssätt	15
2.5	Dekomponering.....	17
2.6	Utvärdering.....	18
2.6.1	Kvalitetskrav.....	18
2.6.2	Praktiska aspekter på utvärderingen	19
3	Implementering.....	20
3.1	Upplägg av nytt system	20
3.2	Ett provsystem och mappstruktur	20
3.2.1	Beskrivning av provsystemet.....	21
3.3	Modellöversyn.....	22
3.3.1	Grundläggande principer	22
3.3.2	Översyn av befintlig modell för säsongrensning	23
3.3.3	Modellöversyn i praktiken	25
3.3.4	Efter modellöversynen	27
3.4	Löpande stöd för säsongrensningen.....	28

4	Bilagor	29
4.1	Återanvändning av Vägledningens provsystem	29
4.2	Modell för säsongrensning och kalenderkorrigering	30
4.2.1	RegArima modell	30
4.2.2	Kommentarer om antal outliers	31
4.3	Kalenderkorrigeringsprinciper	31
4.3.1	PROC X12 anpassning	33
4.4	Andra kalendervariabler: påskeffekter	33
4.4.1	Inbyggda påskeffekter	33
4.4.2	Exempel	34
4.4.3	Egen påskvariabel	35
4.4.4	Andra kalendereffekter: skottår och månads-/kvartalslängd ...	35
4.5	Exempel SAS programkod	35
4.6	Utvärdering av säsongrensningmodeller	36
4.6.1	Jämförelse mellan preliminärt valda modeller	36
4.6.2	Säsongdekomponeringen	37
4.6.3	Andra kriterier	37
5	Litteraturlista	40

1 Syftet med denna vägledning

Regelbundna observationer av en variabel bildar en tidsserie. Tidsserier uppvisar ofta periodiska variationer, normalt är variationen kopplad till vilken tidpunkt under året ett värde avser. Beroende på den studerade serien kan variationen vara stor eller liten. Oavsett storlek vållar den potentiellt visst besvär vid jämförelse över tiden och analys av trender eller konjunkturer. Säsongrensning syftar till att förbättra jämförbarheten över tiden, och därmed förbättra möjligheten att använda tidsserier till meningsfulla slutsatser.

X-12-ARIMA (U.S. Census Bureau, 2010) är fastställt standardverktyg för säsongrensning vid SCB. Metoden är utvecklad av U.S. Census Bureau och används av statistikbyråer runt om i världen. Det fastställda standardverktyget vid SCB går närmare bestämt ut på att metoden tillämpas i form av proceduren PROC X12 i programsystemet SAS®. Ett alternativt verktyg som får användas efter särskild dispens är TRAMO-SEATS från Banco de España (2011).

SCB strävar efter god kvalitet i säsongrensade data. För att säkra detta baseras säsongrensningsarbetet på internationellt utvecklade standarder och kvalitetskrav. De principer som beskrivs i *ESS Guidelines for Seasonal Adjustment* (Mazzi G. L., Calizzani C., 2009) eftersträvas i första hand.

Denna vägledning riktar sig till metodpersonal som är involverad i säsongrensningsarbete vid SCB. Den är avsedd som ett praktiskt stöd och beskriver rutiner för säsongrensningsarbetet vid SCB. Vägledningen går igenom hur de grundläggande principerna tillämpas utifrån förutsättningar som råder för arbetet på SCB.

1.1 Två relaterade dokument

Denna vägledning ingår i SCB:s verksamhetsstöd (VST, delprocess 6.1) tillsammans med två relaterade dokument om säsongrensningsmetoder, nämligen:

- i. [Praxis för säsongrensning](#)
- ii. [Orientering om säsongrensning](#)

Praxisdokumentet (i) riktar sig liksom denna vägledning till metodpersonal som medverkar i arbete med säsongrensning. Praxisdokumentet är mer sammanfattande i punktform med betoning av gemensamma arbetssätt, medan denna vägledning i sin tur avses ge mer av förklaringar, motiveringar och exempel. Orienteringsdokumentet (ii) visar översiktligt de stora dragen, och det riktar sig till en bredare krets av personer som har att göra med tidsseriedata, även externa statistikanvändare.

1.2 Från teoretiskt ramverk till praktisk tillämpning

Säsongrensning är ett omfattande teoretiskt område. I säsongrensningslitteraturen finns ett stort antal artiklar där olika delar av teorin behandlas. De problem som kan dyka upp då en eller flera tidsserier ska säsongrensas varierar och det finns normalt inte ett enkelt recept att följa för att komma till rätta med dessa. Det är inte trivialt att ta till sig artiklar inom

säsongrensning och kunna tillämpa innehållet i den egna praktiska verkligheten. Vägledningen avser att koppla det teoretiska ramverket till praktisk tillämpning på SCB.

1.3 Vägledningens upplägg

Denna vägledning består huvudsakligen av två delar. I den första huvuddelen (avsnitt 2) behandlas principer och metoder, och deras tillämpning. Där ges en teoretisk översikt och en beskrivning av praktiska överväganden i säsongrensningens arbete. Där beskrivs bl.a. kalenderkorrigering, outlierhantering och val av ARIMA-modell. Vidare ingår SAS-kod som exemplifierar olika delmoment.

Vägledningens andra huvuddel (avsnitt 3) går närmare in på implementeringen av bearbetningsrutinerna i SCB:s produktionsmiljö. Den inleds med arbetsgången då ett nytt system för säsongrensning läggs upp. Här ingår moment relaterade till X-12-ARIMA samt ansvarsfördelning mellan säsongrensningens ansvarig och ämnesenheten.

Vanligen utvecklas en tidsseries egenskaper över tid, och för att bibehålla god kvalitet i säsongrensade data behöver säsongrensningen ses över med jämna mellanrum. I vägledningen beskrivs hur detta genomförs i praktiken (se avsnitt 3.3).

2 Principer och metoder

Säsongrensning omfattar ett antal steg. I de olika stegen görs avvägningar med hänsyn till ämneskunskap relaterad till den enskilda serien samt kunskap om säsongrensningens metodernas teori och konventioner. Principer för outlierhantering, kalenderkorrigering och modellval stödjer arbetet, och ger en grund för förtroende för de säsongrensade serier som SCB redovisar. Avvägningarna görs med sikte på en säsongrensad serie med god användbarhet och statistisk kvalitet.

Arbetet med att lägga upp ett system för säsongrensning omfattar dels förberedande frågor att ta hänsyn till, dels praktisk tillämpning av verktyget X-12-ARIMA. De förberedande frågorna omfattar ämnesrelaterade frågor som ska diskuteras med ämnesenheten. Här ingår frågor om konsistens, kalenderkorrigering och outlierhantering.

I praktiken består säsongrensning med X-12-ARIMA (PROC X12 i SAS) av två delprocedurer som i sin tur innehåller flera steg beroende på valda modellinställningar. *Den första delproceduren* syftar till att renodla serien genom att anpassa en s.k. RegArima modell (*regression with Arima-errors*) som illustreras i avsnitt 4.2.1. Se också SAS-programkod i avsnitt 4.5. Tidserien som ska säsongrensas är beroende variabel i regressionsmodellen. Förklaringsvariabler i modellen speglar effekter av sådant som kalendern, outliers, interventionsvariabler eller annat som kan tänkas störa skattningen av säsongkomponenterna. (Sådana effekter kallas ibland något oegentligt för "icke-linjära" effekter.) Därför är specifikationen av RegArima-modellen avgörande för utfallet av säsongrensningen. Säsongrensningsspecialister koncentrerar sig vanligtvis på den här delen genom att testa olika modellalternativ och väljer en modell som uppvisar bästa egenskaper utifrån valda kriterier.

I den andra delproceduren görs en *dekomponering* av den "renodlade" serien från den första delproceduren. Dekomponeringen resulterar i skattningar av tre huvudkomponenter: säsongkomponent, trend-cykel och irreguljär komponent. Detta sker i programvaran med hjälp av särskilda vägda glidande medelvärden, kallade *filter*. Den här delen kan i någon mån styras genom styrparametrar för valet av filter.

I texten som följer förklaras de viktigaste begreppen relaterade till modellspecifikationer som normalt ingår i säsongrensning på SCB.

2.1 Förberedelser

2.1.1 Allmänt

Initiativ till att sätta upp ett system för säsongrensning kommer normalt från förvaltningsobjektet för produkten eller från ett utvecklingsprojekt, se dokumentet [Praxis för säsongrensning](#). Arbetet med det nya systemet utförs i samarbete mellan ämnes- och metodansvariga, när så är lämpligt gärna delvis i samråd med kvalificerade statistikanvändare.

Det finns ett antal frågor att ta hänsyn till, struktur för säsongrensningssystemet, konsistensproblematik och frågor kring kalenderkorrigering och outlierhantering. I texten som följer berörs struktur för säsongrensningssystemet och konsistensfrågor. Kalenderkorrigering och outlierhantering beskrivs i separata avsnitt längre fram.

Förvaltning och utveckling av system för säsongrensning förenklas om de bygger på en gemensam struktur. Att strukturera och dokumentera system på ett genomgående likartat sätt bidrar till att den gemensamma arbetsgången blir tydlig. Ett exempel på struktur för ett säsongrensningssystem visas i avsnitt 3.2.

2.1.2 Konsistensfrågor

När ett system av tidsserier säsongrensas finns det risk att den konsistens ursprungsserier uppvisar sinsemellan bryts när de enskilda serierna säsongrensas. Den konsistens ursprungsserierna uppvisar kan brytas på flera sätt.

Ett exempel är försörjningsbalansen inom nationalräkenskaperna där BNP byggs upp av ett antal underserier. Ursprungsserierna är konsistenta med BNP genom att BNP består av summan av de underliggande delarna. Vid säsongrensning bryts den naturliga summakonsistensen.

En lösning är att använda det *summakonsistensverktyg* SCB utvecklat (SCB, 2009-2010). Verktöget fördelar ut diskrepansen mellan säsongrensade BNP och de säsongrensade delarna, så att summan av de underliggande delarna åter blir lika med BNP. Först säsongrensas BNP och delarna direkt, och därefter justeras delarna så att summakonsistens uppnås. Det här tillvägagångssättet kallas för *rekonciliering*.

Rekoncilieringen utgår alltså från en direkt ansats för säsongrensning. Det är dock möjligt att kombinera rekoncilieringen med inslag av indirekt ansats

om det finns fler än en summeringsnivå. Detta sker i implementeringen av summakonsistensverktyget i BNP kvartalsdata.

Ett annat alternativ är att använda *indirekt säsongrensning*. Vid indirekt säsongrensning utförs säsongrensning på delarna som bygger upp ett aggregat, och sedan summeras delarna ihop.

Vilken lösning som är lämpligast beror på de underliggande seriernas egenskaper. Val av ansats görs i samarbete mellan ämnes- och metodpersonal, och lämpligen även externa användare, i en metodöversyn.

På SCB föredras vanligen direkt säsongrensning framför indirekt men undantag kan förekomma (se Praxis för säsongrensning, avsnitt 2.1.1.3). Indirekt ansats är t.ex. berättigad när delserier håller god kvalitet och har varierande statistiska egenskaper (t.ex. avvikande säsongmönster mellan olika delserier). Ytterligare ett fall där det kan vara lämpligt med indirekt ansats är när aggregat och delserier är relativt enkelt relaterade till varandra, exempelvis när aggregatet fås som summan av två stora delar.

Skälet att rekommendera direkt ansats ligger i att viktiga totalserier inom SCB:s system oftast håller relativt bra kvalitet vilket inte alltid är fallet med serier på mer förfinade nivåer. Säsongrensning för delserier som innehåller mycket brus (variation som inte förklaras av modellen) i förhållande till variationen som förklaras av modellen brukar leda till osäkra skattningar. Vanligen reduceras bruset i och med aggregeringsprocessen före säsongrensningen vilket eliminerar en del av osäkerheten i stora totalserier.

Ytterligare en konsistensfråga rör serier där det inte är fråga om summakonsistens, men där det är önskvärt att den *utveckling* (i procentuella förändringstal) som de säsongrensade serierna visar inte blir logiskt motstridig mellan olika aggregeringsnivåer. Det finns inte några generellt optimala lösningar för konsistensproblem av sådant slag. I praktiken brukar inkonsistenserna minska om liknande modellinställningar implementeras för relaterade serier. Det kan vara svårt att åstadkomma i stora system så fokus borde ligga på att isolera ett fåtal viktiga delserier och försöka anpassa likartade modellinställningar för de delserier som har störst påverkan på totalserien. Särskilt viktigt är att försöka anpassa hanteringen av outliers för relaterade serier genom att t.ex. låsa typ och position av outliers på ett samordnat sätt. Mer om detta i rapporter från två projekt för outlierhantering (SCB, 2010-2011; SCB, 2014).

I en del tillämpningar kan det vara önskvärt att behålla konsistens mellan kvartalsvärden och årsvärden. De säsongrensade kvartalsvärdenas summer över åren ska då stämma med redovisade årsvärden.

I säsongrensningssammanhanget kallas metoder och rutiner som åtgärdar den här typen av inkonsistens för *benchmarking*. I PROC X12 finns ett kommando för att uppnå konsistens mellan kvartal och år (kommando "X11/Force= ...", se SAS Institute Inc, 2012, s. 2624). Denna rutin ska tas till endast försiktigt, med kontroll av att den inte ger upphov till konstlade effekter såsom kraftiga ändringar i utvecklingstal eller liknande. Benchmarking rekommenderas inte om kalenderkorrigering med handelsdagseffekter (eng. *trading day effects*) implementeras inom säsongrensningssmodellen. Dessutom är benchmarkingrutinen känslig för

förändringar i säsongmönster och ska undvikas om detta kan väntas förekomma.

Starka användarkrav eller bedömningar från ämnessidan ska vara vägledande för beslut om eventuell implementering av benchmarking. Dock kan ankomma på metodpersonal att informera produktansvariga om eventuella konstlade effekter som benchmarking kan föra med sig.

På SCB har ytterligare några metoder för benchmarking förekommit, t.ex. MinD4 (SCB, 2008). Rutinerna i fråga är skräddarsydda för SCB:s behov men bygger på *Dentons metod* som också ligger till grund för den inbyggda rutinen i PROC X12. På SCB används huvudsakligen de här metoderna för avstämning mellan orensade serier på års- resp. kvartals- eller månadsbasis. Detta är användbart för årsserier och motsvarande kvartals- eller månadsserier som kommer från olika källor, eller mäts på olika sätt, men avser samma målstorheter.

Rutinen för benchmarking i PROC X12 rekommenderas under normala omständigheter när det finns uttalat behov av överensstämmelse mellan årliga summor av orensade värden och motsvarande årliga summor av säsongrensade kvartals- eller månadsvärden. Rutinen är så utformad att den åstadkommer grundvillkoret samtidigt som den försöker bevara konsistensen mellan säsongrensade förändringstal och förändringstal för motsvarande värden som justeras m.h.a. benchmarking, i så stor utsträckning som möjligt.

I exemplet nedan (figur 2.1) illustreras årsanpassning för säsongrensade värden för tre huvudserier inom industriproduktionsindex-systemet.

```

PROC X12  AUXDATA= WORK.KALFAKTOR
          DATA= WORK.IPIserier_201402
          SPAN=('Jan2000', )
          DATE= DATUM
          INTERVAL= Month
          MDLINFOOUT= WORK.Modl_02_IPI
          OUTSTAT= Modl_02_OutStat;
          EVENT TC01JAN2009d /USERTYPE = TC;
/* egen låst outliereffekt */
          VAR SNI_B_C SNI_C SNI_B ;
          TRANSFORM function= Auto;
          CHECK PRINT= all;
          REGRESSION USERVER= Ln_Kalf/ USERTYPE= LOM;
          AutoMdl print = AutoChoiceMdl
                    Best5Model
                    UnitRootTest UnitRootTestMdl ;
          ESTIMATE;
          FORECAST;
          OUTLIER SPAN=('Jan2000', )
                  TYPE=( AO TC LS ) ;
          X11 FORCE= TOTALS;
/* tvingar årliga summor av månadsvärden att stämma överens med
motsv. årsvärden för orensade serier */
          OUTPUT OUT= WORK.IPI_Modl_02_OUTSER
                   A1 A8 B1 D10 D11
                   D11A D12 D13 D18
                   E5 E6 E6A E7 E8
                   ;
RUN;

```

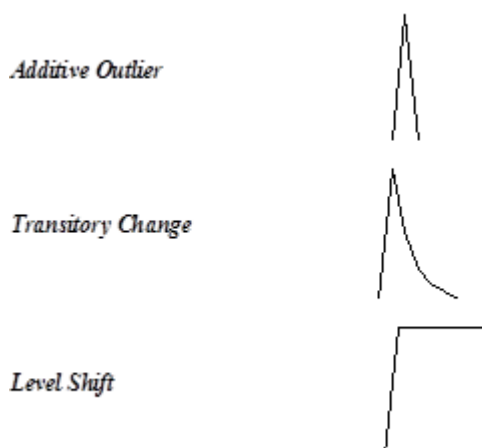

Figur 2.1 Proc X12 exempel med benchmarking (Force= Totals). Säsongrensade värden finns i tabellen D11 medan säsongrensade värden som är korrigerade för årsanpassning finns i tabellen D11A (Output Out=...).

2.2 Outlierhantering

2.2.1 Allmänt

Outlierhantering i säsongrensning handlar om kriterier för när och hur värden i tidsserier ska hanteras som outliers (utstickande värden) och särbehandlas för att inte störa beräkningarna i säsongrensningen. Det finns flera olika typer av outliers, som alla skulle kunna störa och leda till missvisande resultat om de inte togs om hand.

I SAS finns tre fördefinierade outlier typer, *Additive Outlier (AO)*, *Level Shift (LS)* och *Transitory Change (TC)*. En observation (ett värde i tidsserien) klassas som AO om tidsserien just då tillfälligt avviker från sitt normala mönster. I en situation då däremot tidsseriens nivå förskjuts bestående, från och med en tidpunkt, används LS. Om i stället nivån skiftar vid en tidpunkt för att sedan successivt återgå till den tidigare nivån används TC.



Figur 2.2 Outliertyper.

I avsnitt 2.2.2 visas programkod för att testa och justera för de tre outlier typerna.

X-12-ARIMA baseras på vägda glidande medelvärden. Medelvärden påverkas lätt av värden som avviker avsevärt från tidsseriens normala mönster. Om tidsserien inte justeras för outliers kan dessa missvisande påverka skattningen av de komponenter som bygger upp tidsserien och som används då en säsongrensad tidsserie beräknas.

Vid SCB säsongrensas både system av tidsserier med krav på summakonsistens och enskilda tidsserier. I båda fallen är det viktigt med en strategi för outlierhantering för att hålla kontroll på inverkan på den slutliga säsongrensade serien.

För ett system av tidsserier borde en strategi ta hänsyn till samordning av outliers mellan de ingående serierna- samt till hur outliers ska hanteras

inom respektive serie. En samordnad strategi innebär att outliers som identifieras i en serie får påverka outlierinställningarna för övriga serier. Ett exempel på samordnad strategi är *hård top-down*, vilket innebär att en outlier identifierad i en serie på den högsta nivån i ett hierarkiskt system medför att alla delserier påtvingas hantera motsvarande observation som outlier av samma typ. *Exempel: Kvartal 4 år 2008 identifieras som LS i totalserien för export. Hård top-down innebär att de två delserierna export av varor och export av tjänster också hanterar kvartal 4 år 2008 som LS.*

Då en ny observation (produktionsomgång) läggs till i en tidsserie och säsongrensningssproceduren tillämpas på den förlängda serien, så kommer resultatet från de två säsongrensningssomgångarna skilja sig. En anledning till att den säsongrensade serien förändras är att olika outliers kan ha identifierats i de två säsongrensningssomgångarna. En ytterligare fråga att ta hänsyn till är låsning av de outliers som identifieras. Syftet med låsning är att undvika att outliers hoppar in och hoppar ur mellan säsongrensningssomgångar, dvs. att observationer klassas som outlier i en omgång men inte i nästa. Det här är något som kan leda till stora revideringar från en produktionsomgång till en annan.

En tänkbar strategi för outlierhantering i ett hierarkiskt system av serier med krav på summakonsistens är *kombinationen av hård-top down och mjuk låsning*. *Hård top-down* innebär att när en observation i tidsserien på den högsta nivån klassas som outlier ska outliers införas i underliggande serier vid samma tidpunkt. Med *mjuk låsning* menas att nya outliers, utöver dem som redan identifierats och låsts, tillåts i den period där låsta outliers förekommer (SCB, 2014).

Outlierhantering utförs innan tidsserien dekomponeras. De outliers som säsongrensningssproceduren identifierar ersätts med modellvärden för att skattningen av de olika komponenterna inte ska störas. Då säsongrensningen är genomförd läggs outliers tillbaka i den säsongrensade serien. Tanken är att outliers bär på information kring viktiga händelser, händelser som är intressanta och relevanta när den säsongrensade serien analyseras, och när den tjänar som beslutsunderlag hos användare.

I en ny produktionsomgång ska alltid tidsserien justeras för outliereffekter. Både ämneskunskap och statistisk teori ligger till grund i bedömningen av om en observation ska klassas som outlier eller inte. Bedömningarna i de kvartals- eller månadsvisa produktionsomgångarna följer principer som lagts fast i de årsvisa metodöversynerna, för att inte öppna för förhastade reaktioner. Normalt ska automatisk outlierhantering användas för hela den tidsserie som ska säsongrensas om det inte finns skäl till att låsa vissa historiska outliers. Låsning används för att minska revideringar i enskilda serier eller för att behålla konsistens mellan olika relaterade tidsserier. Låsningen innebär att historiska outlierpositioner blir fasta, d.v.s. typ och position av outliers hålls oförändrade i kommande produktionsomgångar samtidigt som skattningen för outliereffekten alltid utförs. Låsningen kan antingen göras för samtliga observationer i tidsserien (förutom för den sista observationen) eller för en delmängd m.h.a. outliers som blivit identifierade före en viss tidpunkt. Det sistnämnda är mest förekommande angreppssättet då låsningen vanligtvis görs för den delen av serien som inte påverkas av revideringar i orensade data. I praktiken låser man typ och position av

outliers för alla observationer utom för de senaste 4-5 observationerna, beroende på hur långt tillbaka i tiden revideringen görs för serien i fråga.

2.2.2 Exempel med outlierhantering

I programkoden nedan (figur 2.3) illustreras säsongrensning för tre av huvudserierna för systemet för Industriproduktionsindex (SNI_B_C, SNI_C respektive SNI_B) som finns i tabellen IPISerier_201402.sas7bdat¹.

Modellspecifikationen innehåller outlierkorrigeringskommando `OUTLIER`. Här anges samtliga typer av outliers med kommando `TYPE`.²

Dessutom används kommando `EVENT` för att låsa typ och position av en i förväg bestämd outliereffekt i tidpunkten för januari 2009.

Tilläggskommando `/USERTYPE = TC` medför att programmet tolkar händelsen som en temporary change (TC) och inte som en ordinarie indikatorvariabel (dummy). I det fallet kommer den skattade effekten att tilldelas den irreguljära komponenten som är en del av den slutliga säsongrensade serien.

Programmet kommer att fortsätta söka efter andra outliertyper i alla tidpunkter i perioden mellan januari 2000 och november 2008, tack vare specifikationen `OUTLIER` och kommandot `SPAN=(' Jan2000' , ' Nov2008')`. I det här fallet kommer en enda outliereffekt att identifieras efter november 2008 och det är den låsta outliereffekten i januari 2009. Skulle man vilja att programmet identifierar eventuella ytterligare outliers i samtliga tidpunkter skulle det räcka med att utesluta `' Nov2008'` och behålla `' Jan2000'` eftersom januari 2000 är startdatum för serierna.

Programmet kommer att justera tidsseriens komponenter med de skattade effekterna, beroende på typerna av identifierade outliers.

Trendkomponenten justeras med LS (level shift) medan den irreguljära komponenten justeras med de skattade TC- och AO-effekterna.

¹ Finns här: P:\Prod\PCA\MFFM_Tidseriegrupp\TidserieGrupp\IPI-Mall\Data\Indata.

² Default i PROC X12 är outliers av typ AO och LS (om man inte specificerar `TYPE = ...`).

```
PROC X12 AUXDATA= WORK.KALFAKTOR
DATA= WORK.IPIserier_201402
SPAN=('Jan2000', )
DATE= DATUM
INTERVAL= Month
MDLINFOOUT= WORK.Modl_03_IPI
OUTSTAT= Modl_01_OutStat;
EVENT TC01JAN2009d /USERTYPE = TC;
VAR SNI_B_C SNI_C SNI_B ;
TRANSFORM function= Auto;
CHECK PRINT= all;
REGRESSION USERVAR= Ln_Kalf/ USERTYPE= LOM;
AutoMdl
print = AutoChoiceMdl
      Best5Model
      UnitRootTest UnitRootTestMdl ;
ESTIMATE;
FORECAST;
OUTLIER SPAN=('Jan2000', 'Nov2008')
        TYPE=( AO TC LS ) ;

X11;
OUTPUT OUT= WORK.IPI_Modl_03_OUTSER
          A1 A8 B1 D10 D11
          D12 D13 D18 E6
          ;

RUN;
```

Figur 2.3 PROC X12 exempel med outlierhantering.

2.3 Kalenderkorrigering

Kalenderkorrigering ska normalt implementeras om de kalendermässiga variationerna i arbetsdagarnas (eller motsvarande) fördelning över året kan bedömas vara märkbart bidragande till seriens variationer. Både ämnesmässiga och metodologiska skäl kan ligga bakom bedömningen om eventuell implementering. Korrigering för kalendereffekter kan ha särskilt synlig betydelse när värdet för en månad resp. ett kvartal jämförs med värdet för motsvarande månad eller kvartal föregående år. Ett tydligt skäl för kalenderkorrigeringen är därför att förbättra noggrannheten i skattningar för årsutveckling i tidserier där kalendereffekter inverkar.

En del konsekvenser av kalendern behandlas inte som "kalendereffekt" utan tillåts slå igenom som säsongeffekt, då de ingår i den periodiska delen av årlig naturlig variation. T.ex. månadslängden får påverka säsongvariationen eftersom antalet dagar i en månad är detsamma oavsett vilket år det handlar om. Ett undantag är februari under skottår, vilket innebär att skottåreffekten kan behöva behandlas som en kalendereffekt.

Vidare varierar antalet arbetsdagar i en månad från år till år beroende på när t.ex. helgdagar infaller, vilket ger en kalendereffekt. Särskilt påsken infaller på olika dagar olika år vilket gör att *påskeffekten* potentiellt kommer in som en kalendereffekt. Andra kalendereffekter kan förekomma för speciella typer av data, exempelvis handelsdagseffekter (eng. *trading day*) vilket är typiskt för månadsstatistik från t.ex. detaljhandeln. Idealt kunde vara att korrigera tidsserier genom att ta hänsyn till varierande uppdelningar av veckodagar och helger i de månader man jämför med (se avsnitt 4.2.3 för mer detaljer kring påsk- och handelsdagseffekter).

SCB har som huvudprincip att implementera kalenderkorrigering som en del av den ordinarie säsongrensningprocessen. Detta görs genom att

kalendervariabler skapas i förväg och används som förklaringsvariabler i förjusteringsdelen som ingår i RegArima-proceduren (delprocedur i X-12-Arima via PROC X12 i SAS).

Kalendervariablerna skapas med hjälp av svensk kalender. Normalt används en variabel som beskriver hur antalet arbetsdagar i den aktuella perioden avviker från det genomsnittliga antalet arbetsdagar för motsvarande månad eller kvartal på lång sikt.

Att på det sättet utgå från antal arbetsdagar enligt den gängse kalendern kan vara acceptabelt som en konvention och medveten förenkling. Teoretiskt kunde det annars vara närmare ett ideal att välja kalendervariabeln efter bransch, såsom "handelsdagar" i någon mening för en del branscher. Att i stället genomgående utgå från arbetsdagar gör å andra sidan bearbetningen enhetligare, vilket beroende på användarnas förväntningar kan vara till fördel på sitt sätt för resultatens tolkbarhet.

Exemplet i figur 2.4 illustrerar SAS-kod där två typer av kalendereffekter används, en egen endogen i form av förklaringsvariabeln `ln_kalf` och en inbyggd (EASTER) som hanterar påskeffekt.

Detaljerat tillvägagångssätt för beräkningen av kalenderfaktorer redovisas i avsnitt 4.2.2.

```
Title 'PROC X12 med arbetsdagseffekt och påsk' ;
PROC X12 AUXDATA= WORK.KALFAKTOR
DATA= WORK.IPIserier_201402
SPAN=('Jan2000', )
DATE= DATUM
INTERVAL= Month
MDLINFOOUT= WORK.Modl_04_IPI
OUTSTAT= Modl_04_OutStat;
***EVENT TC01JAN2009d /USERTYPE = TC;
/* egen låst outliereffekt */
VAR SNI_B_C SNI_C SNI_B ;
TRANSFORM function= Auto;
CHECK PRINT= all;
REGRESSION PREDEFINED= EASTER(7) ;
/* Inbyggda kalenderfaktorer: Påsk(EASTER) */
REGRESSION USERVAR= Ln_kalf/ USERTYPE= LOM ;
/* egen kalenderfaktor 'Ln_kalf' med motsvarande
förutsättningar som de inbyggda faktorerna
(specificeras med LOM) */
AutoMdl print = AutoChoiceMdl
Best5Model
UnitRootTest UnitRootTestMdl ;
ESTIMATE;
FORECAST;
OUTLIER SPAN=('Jan2000', )
TYPE=( AO TC LS ) ;
X11 ;
OUTPUT OUT= WORK.IPI_Modl_04_OUTSER
A1 A8 B1 D10 D11
D11A D12 D13 D18
E5 E6 E6A E7 E8
;
RUN;
```

Figur 2.4 PROC X12 exempel med en egen och en inbyggd kalenderfaktor.

2.4 Val av ARIMA-modell

2.4.1 Allmänna hållpunkter

När outlier- och kalendereffekter skattas använder X-12-ARIMA en så kallad RegARIMA-modell. Den observerade tidsserien antas kunna beskrivas av en regressionsmodell bestående av outlier- och kalendereffekter där residualtermens slumpmekanism antas följa en ARIMA-modell (se avsnitt 4.2.1).

Valet (specificeringen) av RegARIMA-modellen är ett av leden i säsongrensningssystemet.

Då ett nytt säsongrensningssystem byggs upp finns större spelrum i valet av RegARIMA-modell. Ofta kan det vara möjligt att välja mellan flera olika

ARIMA-modeller, vägledande i valet är modellens statistiska egenskaper. Modellval ska göras med hänsyn till:

- Signifikanta parameterskattningar – om parameterskattningar inte är signifikanta bör modellinställningarna ses över. Det här gäller först och främst för parametrar som avser egna förklaringsvariabler, t.ex. kalenderfaktorer. Signifikans för parametrar som definierar val av ARIMA-modell (AR, MA, differentieringsordning) utvärderas när flera alternativa modeller testas på ett manuellt sätt.
- Antal skattade parametrar, en modell med färre skattade parametrar är allmänt sett att föredra framför en med många (AR- och MA-) parametrar.
Notera att i samband med automatisk modellsökning med specifikationen AUTOMDL i PROC X12 väljer programmet (automatiskt) modeller som har färre parametrar framför modeller med fler parametrar utifrån inbyggda kriterier.
- Differentieringsordning: en modell med låg differentieringsgrad inom ARIMA-specifikationen för ordinarie- och säsongdifferens föredras.
- Tester för RegArima- residualer:
 - Normalitet: utvärdering av snedhet (skewness), toppighet (kurtosis).
 - Autokorrelationer: En modell med markant autokorrelation tyder på dålig anpassning. Här kan Ljung-Box test med fördel användas för att exempelvis jämföra två konkurrerande modeller.
 - Kontrollera för eventuell kvarvarande säsongvariation: Utstickande autokorrelationer gällande lag 12, 24, 36 för månadsdata respektive lag 4, 8, 12, 16, osv för kvartalsdata.
 - Tester i spektraldomän som t.ex. Bartlett's Kolmogorov-Smirnov test i proceduren PROC SPECTRA (ingår ej i PROC X12).
 - Avvikelse från linjaritet i modellen, enligt residualplottar.
- Informationskriterier (AICC eller BIC) för att välja bland flera jämförbara modeller.
- Outliers: Antal identifierade outliers (om automatisk outlierhantering görs), en modell med färre upptäckta outliers föredras framför en med alldeles för många outliers. Koppling till verklighet: Modellens anpassningsförmåga att fånga viktiga händelser som kan relateras till serien ifråga genom att hantera vissa observationer som outliers.(se avsnitt 4.2.2).
- Prognosförmåga (t.ex. average forecast error).
- Revideringsbenägenhet – en modell som orsakar svårförklarliga och/eller stora revideringar i säsongrensade värden mellan två på varandra följande omgångar anses inte uppfylla grundläggande kvalitetskriterier.

Arbetsgången i modellvalet blir delvis olika då man sätter upp ett nytt säsongrensningssystem mot då man ser över ett befintligt system. Principer för modellöversyn går igenom i avsnitt 3.3.

Vid modellöversyn tillämpar SCB s.k. *partial concurrent*-principen (Mazzi G. L., Calizzani C., 2009, s. 23). Detta innebär att fasta modellinställningar används mellan två årliga modellöversyner. Modellparametrarna ska däremot normalt skattas på nytt för varje produktionsomgång (månad resp. kvartal).

Allmänt ska en konservativ ansats tillämpas vid modellöversyn. Detta innebär att modellinställningar endast ska ändras om det är motiverat att inte behålla gamla modellinställningar, och i så fall vid planerad modellöversyn, såvida inte synnerliga skäl talar för annat.

2.4.2 Tillvägagångssätt

Här ges en grundläggande beskrivning steg för steg av en tänkbar arbetsprocess för modellval. Se figur 2.5 för schematisk beskrivning av en tänkbar arbetsgång. Detaljerna kan variera från fall till fall beroende på omständigheterna och eventuella specifika krav inom varje system.

Om systemet är stort kan det vara lämpligt att utgå från automatisk modellidentifiering, den automatiska identifieringen resulterar i ett modellförslag för varje serie utifrån ett antal inbyggda kriterier. Denna metod kallas här för alternativ A.

Alternativ B å andra sidan innebär att en grundmodell, vanligen en s.k. Airline-modell (ARIMA (0,1,1)x(0,1,1)), används för alla serier. Det här alternativet kan vara lämpligt för ett system med få tidserier. Justeringar av grundmodellen görs enbart för de serier som visar dåliga diagnostiska resultat, när grundmodellen provas på dem.

Kombination av alternativ A och B, i någon form, kan vara ett tredje alternativ. Exempelvis skulle man kunna tänka sig att använda automatisk modellsökning för en grupp av tidserier samtidigt som modellering för resten av tidserierna skulle utgå ifrån Airline-modellen.

2.4.2.1 Alternativ A: utgå ifrån automatisk outlierhantering

Automatisk modellsökning utförs genom kommandot `AUTOMDL`. Se ETS User's Guide för förklaringar kring syntaxen (SAS Institute Inc., 2012). X12-ARIMA kommer då att föreslå lämpliga modellalternativ som också kan sparas i form av en tabell. Programmet kommer att välja en modell utifrån ett antal förprogrammerade kriterier. Proceduren är baserad på principer från TRAMO-programmet (Time Series Regression with ARIMA Noise, Missing Observations and Outliers, (Maravall, 2011)). Om alternativet `BALANCED` väljs som komplettering för `AUTOMDL` blir RegARIMA i PROC X12 identisk TRAMO. Programmet använder likelihoodbaserade informationskriterier (AIC, AICC, HQ och BIC) för att välja optimala modellinställningar. Där kan ingå val av transformation, eventuella förklaringsvariabler (regressorer), och potentiella outliereffekter.

Att se upp med är att automatiskt modellval kan peka ut modeller som är olika men har kriteriemåttvärden nära varandra. Ifall det finns särskilda

skäl kan den modell som har bättre *övriga* egenskaper väljas istället för den automatiskt valda modellen.

Arbetsgången har normalt följande upplägg:

- Steg 1: Gör en automatisk säsongrensning med frånkopplad outlierhantering och utan kalenderfaktorer.
- Steg 2a: Gör en automatisk säsongrensning genom att lägga till outlierhantering (också automatisk). Notera om modelländringar förekommer.
- Steg 2b: Gör en automatisk säsongrensning genom att lägga till både outlierhantering och kalendereffekter, om krav på kalenderkorrigering förekommer. Kalenderfaktorer väljs i samråd med ämneskunniga (antingen inbyggda eller egna kalenderfaktorer). Notera om modelländringar förekommer.
- Steg 4: Utvärdera output – diagnostiska tester, grafisk analys, ämnesmässiga utvärderingar m.m. För det här steget finns ytterligare information i avsnitt 2.6.
- Steg 5: Gör ändringar i automatiska modellinställningar för de serier där behovet uppstår. Se figur 2.5 för detaljer.
- Steg 6: Spara de nyskapade modellinställningarna som permanenta.
- Steg 7: Utför ytterligare en körning med fasta modellinställningar och utvärdera resultat. Jämför med tidigare omgångar vid behov.

2.4.2.2 Alternativ B

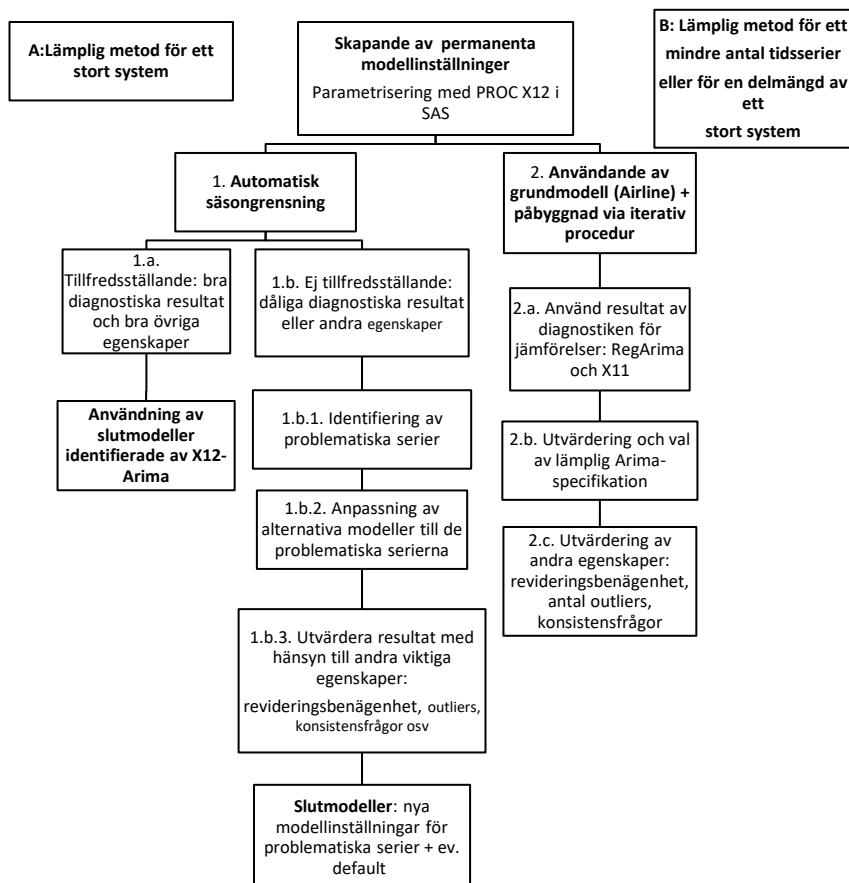
Ifall modellvalet blir försvårat p.g.a. andra egenskaper (t.ex. brus i serien starkt jämfört med signalen) så att programmet inte lyckas hitta en tillfredsställande modell (t.ex. instabilitet över tiden, stora revideringar m.m.) får man utgå från en och samma grundmodell för samtliga serier.

Om förutsättningar tillåter kan Airline-modell (ARIMA(0,1,1)*(0,1,1)) väljas som ett lämpligt alternativ. Den här modellen har visat sig fungera väl för de flesta makroekonomiska tidserier. Den används som referens-modell i samband med automatisk modellsökning i PROC X12. Ett annat alternativ är använda någon annan lämplig grundmodell som resulterar i bra anpassningar för de viktigaste serierna.

Ytterligare ett alternativ som kan vara tidskrävande är att manuellt anpassa flera alternativa modeller för varje serie och utvärdera utfallet. Det här alternativet kan vara lämpligt för några serier som har utpräglad betydelse såsom särskilt viktiga samhällsekonomiska indikatorer.

För illustration av båda alternativen se SAS-programmen:

- Vägledning_del1_AutomModell.sas,
- Vägledning_del2_ModellVal_ProblemSer.sas³



Figur 2.5. Schematisk beskrivning av arbetsprocessen: Skapande av permanenta modellinställningar för säsongrensning med X12-ARIMA.

2.5 Dekomponering

Säsongrensning bygger på idén att den observerade tidsserien (nedan betecknad Y) består av ett antal icke observerbara komponenter. Det är dessa icke observerbara komponenter säsongrensingsproceduren skattar och som ligger till grund för beräkning av den slutliga säsongrensade serien.

³P:\Prod\PCA\MFFM_Tidseriegrupp\TidserieGrupp\Förvaltning\SASkod

Vanligen antas att serien kan delas upp i komponenterna trendcykel (TC), säsongeffekt (S), kalendereffekt (K), irreguljär effekt (I) och outliereffekt (O). Beräkningen av komponenterna sker stegvis. I ett första steg beräknas kalender- och outliereffekter, därefter beräknas övriga komponenter. En *additiv* modell kan skrivas i följande form:

$$Y = TC + S + K + I + O \quad (2.1)$$

För varje tidpunkt antas att den observerade seriens värde Y genererats som summan till höger om likhetstecknet. Den summan består av värdena vid tidpunkten på de nyss nämnda icke observerbara komponenterna.

Ett alternativ till den additiva modellen är en *multiplikativ* modell. I en multiplikativ modell består högerledet i stället av en produkt av komponenterna.

Säsongrensningsverktyget hanterar både additiva och multiplikativa modeller. I någon mening kan det gå att testa vilkendera formen som ger bäst anpassning till serien, men valet mellan additiv och multiplikativ modell har även andra aspekter, som kan vara viktigare. Konsistensskäl talar för att inte i onödan blanda de två formerna inom ett system av relaterade serier. Multiplikativ modell är ofta lämplig för ekonomiska data, men den förutsätter att alla värden i serien är större än noll. För serier med nettoposter som kan gå under noll är bara additiv modell möjlig.

När den observerade tidsserien korrigerats för kalender- och outliereffekter beräknas de resterande komponenterna, S, TC och I, de komponenter som tidigare antogs kunna beskrivas med en ARIMA-modell.

I X-12-ARIMA beräknas dessa komponenter stegvis, i varje steg tillämpas vägda glidande medelvärden (filter). När alla komponenterna beräknats kan den slutliga säsongrensade serien skapas. När den observerade tidsserien antas följa en additiv modell görs det enligt:

$$SR = Y - S - K \quad (2.2)$$

Här står SR för den slutliga säsongrensade serien. Den slutliga säsongrensade serien innehåller alltså outliereffekterna och den irreguljära komponenten.

2.6 Utvärdering

2.6.1 Kvalitetskrav

Kvaliteten i hur väl säsongrensningen fungerar för givna serier kan belysas genom inbyggda teoretiska eller empiriska kvalitetindikatorer. För statistikprodukter med många serier i ett säsongrensningssystem kan det vara orationellt att granska enskilda serier efter varje produktionsomgång. Granskning ska dock alltid göras vid modellöversyner och vid nyutformning av ARIMA-modeller för nya serier. Här nedan anges några viktiga kriterier för utvärdering.

1. Utvärdering av RegArima-delen, som inkluderar både inbyggda och egna kriterier överensstämmer till stor del med proceduren för val av ARIMA-modell som beskrivs i avsnitt 2.4. Här ingår:
 - a. Parameterskattningar (signifikanta eller ej, relevans m.m.).

- b. Tester för residualer från den skattade RegArima modellen: Autokorrelationer (Ljung-Box test), normalitetstester (skevhet och toppighet), tester för kvarvarande säsongvariation och eventuellt andra indikatorer på modellavvikelse, såsom residualplottar.
 - c. Outliers, antal och typ samt koppling till verklighet.
 - d. Prognosnoggrannhet (t.ex. average forecast error).
 - e. Revideringsbenägenhet i termer av instabila ARIMA-modeller från en produktionsomgång till en annan.
2. Utvärdering av X11-delen som har med dekomponering att göra.
 - a. Analys av komponenter: Finns signifikant kvarvarande säsongvariation i den irreguljära komponenten? Är den irreguljära komponenten fri från systematisk variation? Är säsongkomponenten tillräckligt stabil (om *moving seasonality* förekommer).
 - b. Sammanställning av kvalitetsmått (M1-M11, samt översiktliga Q och Q2, U.S. Census Bureau, 2011; SAS Institute Inc, 2012).
 3. Generell översiktlig utvärdering: Kan revideringar i den säsongrensade serien uppfattas som motiverade? Jämförelser mellan två på varandra följande produktionsomgångar, jämförelser mellan årsutvecklingstal, granskning av inkonsistenser mellan relaterade serier m.m. Finns kopplingar med serier från andra produkter och finns skäl att stämma av resultaten?

2.6.2 Praktiska aspekter på utvärderingen

I praktiken kan utvärderingen delvis se olika ut beroende på tid, resurser, typ av system, användarkrav, tidigare erfarenhet om serierna och andra faktorer. Fundamentalt är att såvitt möjligt försäkra sig mot misstag och tekniska störningar i exekveringen av programmet, vilket indikativt kan följas upp i sparade SAS-log-filer. Därefter utvärderas kvaliteten i säsongrensningen, vilket den följande genomgången fokuserar på.

Gör *grafiska utvärderingar* för att skaffa en snabb uppfattning om kvaliteten i säsongrensningen. Särskilt gäller det om många tidserier måste granskas med kort varsel. Grafer av relevanta tidseriekomponenter kan avslöja om det har förekommit problem med t.ex. revideringar eller avvikelser från ett mer eller mindre normalt beteende.

Ett lämpligt första steg i utvärderingsprocessen kan vara att använda ett SAS-makro som gör grafiska jämförelser mellan olika komponenter för varje serie separat, jämförelser mellan två relaterade (t.ex. på varandra följande) produktionsomgångar och liknande utvärderingar av resultat. Se exempel i SAS-programmet JMF_Utvecklingstal_IPI.sas⁴.

I avsnitt 4.6, ges en mer detaljerad beskrivning av utvärderingen av säsongrensningens modeller.

⁴ P:\Prod\PCA\MFFM_Tidseriegrupp\TidserieGrupp\IPI-Mall\Kontroll

3 Implementering

Införande av ett nytt system för säsongrensning i en produkt eller i en grupp av relaterade produkter initieras vanligen av förvaltningsteamet för produkten, eller som resultat av ett utvecklingsprojekt. Det finns flera förutsättningar som behöver utredas och säkras innan en säsongrensningsprocedur kan implementeras för ett system. Allmänna principer för säsongrensning finns i dokumentet Praxis för säsongrensning i Verksamhetsstödet.

3.1 Upplägg av nytt system

Upplägg av system styrs av produktspecifika förutsättningar som inkluderar ämnesfrågor, tekniska krav (IT- och metod) och övriga mer eller mindre definierade krav och specifikationer. Upplägg av ett nytt system görs vanligen genom att följa några steg:

1. Undersök egenskaperna hos serier som ska säsongrenas utifrån ämnesperspektiv och beakta eventuella önskemål (krav) från användarna. Undersök om serierna uppfyller grundläggande tekniska (metodologiska) förutsättningar (SAS Institute Inc, 2012).
2. Återkoppla till ämnessidan (produktansvarig) med information om vad de metodologiska studierna visat om förutsättningarna för säsongrensning av serierna i fråga. Börja planera tillsammans med ämnessidan för metainformation till användare i statistikredovisningen, om eventuella särskilda kvalitetsaspekter såsom konsistensproblem eller oregelbundna serier.
3. Gör en preliminär skiss av en systemstruktur som ska ta hänsyn till specifika krav och önskemål, och formulera tillsammans med ämnessidan hur en tänkbar output ska se ut.
4. Definiera hur den metodologiska sidan av systemet ska se ut. Här ingår planering för modellinställningar som styr säsongrensningsproceduren, och för diagnostiska tester och tabeller som ska sparas. Vidare ingår metodologisk beskrivning av proceduren och liknande.
5. Skapa en grundläggande mappstruktur för säsongrensningssystemet med hänsyn tagen till ämnes-, metodologiska och empiriska krav ovan. Se ett provsystem och beskrivning nedan i avsnitt 3.2.
6. Undersök och specificera var orensade serier finns och granska deras sammansättning (typ av tabell, format m.m.). Specificera var parameterfil och kalendervariabel finns.

3.2 Ett provsystem och mappstruktur

Ett exempel på produktionssystem har skapats, exemplet baseras på Industriproduktionsindex (IPI) men med ett mindre antal serier (10 serier) och en något förenklad mappstruktur. Systemet kan användas som en mall till framtagande av nya system, provsystemet återfinns under:

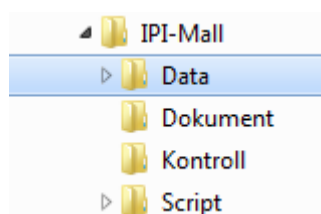
P:\Prod\PCA\MFFM_Tidseriegrupp\TidsserieGrupp\IPI-Mall.

I undermappen *Dokument* finns instruktioner till hur programmen i provsystemet exekveras. En översiktlig beskrivning följer nedan.

3.2.1 Beskrivning av provsystemet

IPI- och provsystemet har konstruerats så att produktionen, i möjlig mån, kan automatiseras. Program och data ligger sparade i separata mappar, även produktionsomgångar sparas separat så att dessa vid behov kan köras om med hjälp av sparad input och programkod.

Provsystemets huvudkataloger består av *Data*, *Dokument*, *Kontroll*, och *Script*, se Figur 3.1. I mappen *Dokument* sparas instruktionen till att köra systemet. Mappen *Kontroll* innehåller program som granskar och utvärderar säsongrensningen. Där finns nu ett SAS-program som jämför utvecklingstal av säsongrensade data och trendskattningar mellan två produktionsomgångar.



Figur 3.1 Huvud- underkataloger för provsystemet.

Mappen *Script* innehåller SAS program, ett huvudprogram som i sin tur anropar andra delprogram som läser indata, genomför bearbetningar, säsongrensar, och genomför efterarbete. Systemet är automatiskt och tekniskt sett behöver normalt inga ändringar göras (se dock instruktionen och beskrivning nedan angående "Prodomgang.txt").

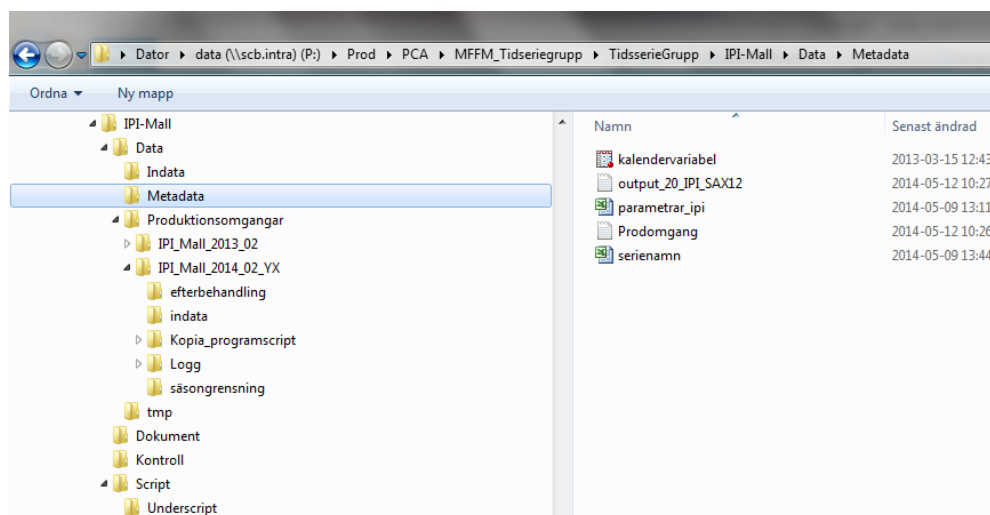
Mappen *Data* innehåller undermapparna *Indata*, *Metadata*, *Produktionsomgångar*, och *tmp*. Indata till säsongrensning sparas i *Indata*. För provsystemet är bara testdata avseende februari 2013 och 2014 tillgängliga.

Mappen *tmp* används av SAS programmet för att skapa tillfälliga data. Innehållet i mappen rensas automatiskt av SAS-programmet.

Metadata innehåller flera viktiga filer som styr inställningar (se höger panel i figur 3.2). Filerna i "metadata" behövs för varje produktionsomgång. T.ex. kalendervariabeln som omarbetats till IPI serier, parameter fil som styr säsongrensning exempelvis ARIMA parametrar, parameter som styr transformation (logaritm eller inte), och outliers.

Filen *serienamn* innehåller serier som ska redovisas som slutresultat. I textfilen *Prodomgang.txt* anges vilken produktionsomgång eller testomgång som säsongrensningen avser (år, månad och tillägg). Baserad på inställningen i textfilen kommer SAS-programmet att skapa en enskild produktionsomgång, genomgå test av indata m.m. För provsystemet är

det endast möjligt att specificera 2013-02 eller 2014-02 på grund av begränsningar i indata.



Figur 3.2 Skärmdokumentet över provsystemet IPI-Mall.

Under mappen *Produktionsomgångar* sparas information om varje utförd säsongrensning. Informationen sparas i mapparna (se vänster panel i Figur 3.2) *indata*, *kopia_programscript*, *logg*, *säsongrensning*, och *efterbehandling*. *Indata* är det aktuella indatamaterialet som används i säsongrensningen, till skillnad från andra mappar (under *Data\indata*) som kommer att ändras över tid. Med sparade indata, parameterfil och aktuell kalendervariabel (under *Säsongrensning*), och *kopia_programscript*, kan varje körning göras om vid ett senare tillfälle efter behov. Loggen om eventuella felindikationer vid avbrutna körningar ligger under mappen *Logg*.

3.3 Modellöversyn

3.3.1 Grundläggande principer

Modellinställningar som innehåller säsongrensningsspecifikationer för ett etablerat system av tidserier ska normalt ses över årligen. Planering för framtida modellöversyner behöver göras i ett tidigt skede redan i upplägget av ett nytt säsongrensningssystem. Planen ska löpande uppdateras i anslutning till förvaltningsplaneringen.

SCB har, som också påpekats i avsnitt 2.4.1, huvudsakligen implementerat en ansats för modellöversyn som på engelska benämns *partial concurrent adjustment*. Ansatsen bygger på att Arima-modeller, inklusive filterspecifikation, outlierspecifikation och eventuella kalender- och andra tänkbara förklaringsvariabler ligger fast mellan två på varandra följande översyner. Däremot görs skattningen av modellparametrarna alltid på nytt i varje ny säsongrensningssomgång (produktionsomgång).

Inställningarna som ligger fast under året ses över och ändras vid behov en gång per år vilket är lämpligt under normala omständigheter. Syftet med den här ansatsen är att möjliggöra att säsongrensningen ska behålla tillräckligt god statistisk kvalitet samtidigt som viktiga användarkrav blir uppfyllda, såsom att revideringar minimeras.

Undantagsvis kan modellöversyn behövas även under löpande år om datamaterialets egenskaper undergår omfattande ändringar så att modellerna borde omprövas. Vanligen görs extra modellöversyner vid större omläggningar i statistiken, t.ex. vid SNI-omläggningar, byte av referensår för ekonomiska tidserier, förlängningar av tidserier, metodbyte i undersökningen och liknande. Detaljer om tider för kommande modellöversyn ska normalt specificeras i förvaltningsplanen.

Mellan två planerade översyner ska modellinställningarna inte ändras om det inte finns synnerliga, exceptionella, skäl för det. Exempel på sådana skäl skulle möjligen kunna vara uppenbart orimliga säsongrensade resultat för viktiga serier, som skulle kunna bero på modellens instabilitet eller felspecificering. Här gäller det att vara synnerligen restriktiv, så att inte statistikens objektivitet äventyras av förhastade reaktioner i akuta lägen.

Eventuella förslag till ändringar kan initieras av metodstatistiker som agerar som stöd för säsongrensning i respektive produkt och/eller ämnesansvariga. De ansvariga metodstatistikerna får undersöka problemen tillsammans med ämnesansvariga och fatta beslut om eventuellt behov av ändring i modellinställningar. Metodstatistiker som har rollen som säsongrensningsspecialister har huvudansvar för utförande av en fullständig och kvalitetssäkrad modellöversyn. Produktansvariga har ansvar att skapa förutsättningar till säsongrensningsspecialister för smidigt utförande av arbetet.

En modellöversyn kräver gedigna teoretiska kunskaper och erfarenhet från praktiskt arbete med säsongrensning. Dessutom är det önskvärt att ha god kännedom om produkten i fråga och om egenskaper hos serier som ska säsongrenas. Vid svåra överväganden kan ansvariga metodstatistiker begära stöd från säsongrensningsexperten som då får avgöra val av modellinställning. Detta sker när t.ex. införande av modelländringar skulle kunna leda till en förändrad tolkning av statistiken utifrån användarsynvinkel. Information från kvalificerade användare kan utnyttjas som kompletterande underlag för övervägandena när så är lämpligt.

3.3.2 Översyn av befintlig modell för säsongrensning

Förfarandet vid en modellöversyn är i stort sett analogt med identifiering av modeller till nya serier som beskrivs i avsnitt 2.4. En skillnad är att de befintliga modellerna, som använts före modellöversynen, är de modeller som den fortsatta utvärderingen bygger på.

Tillvägagångssättet för en normal modellöversyn är upplagt enligt Diagram 3.1

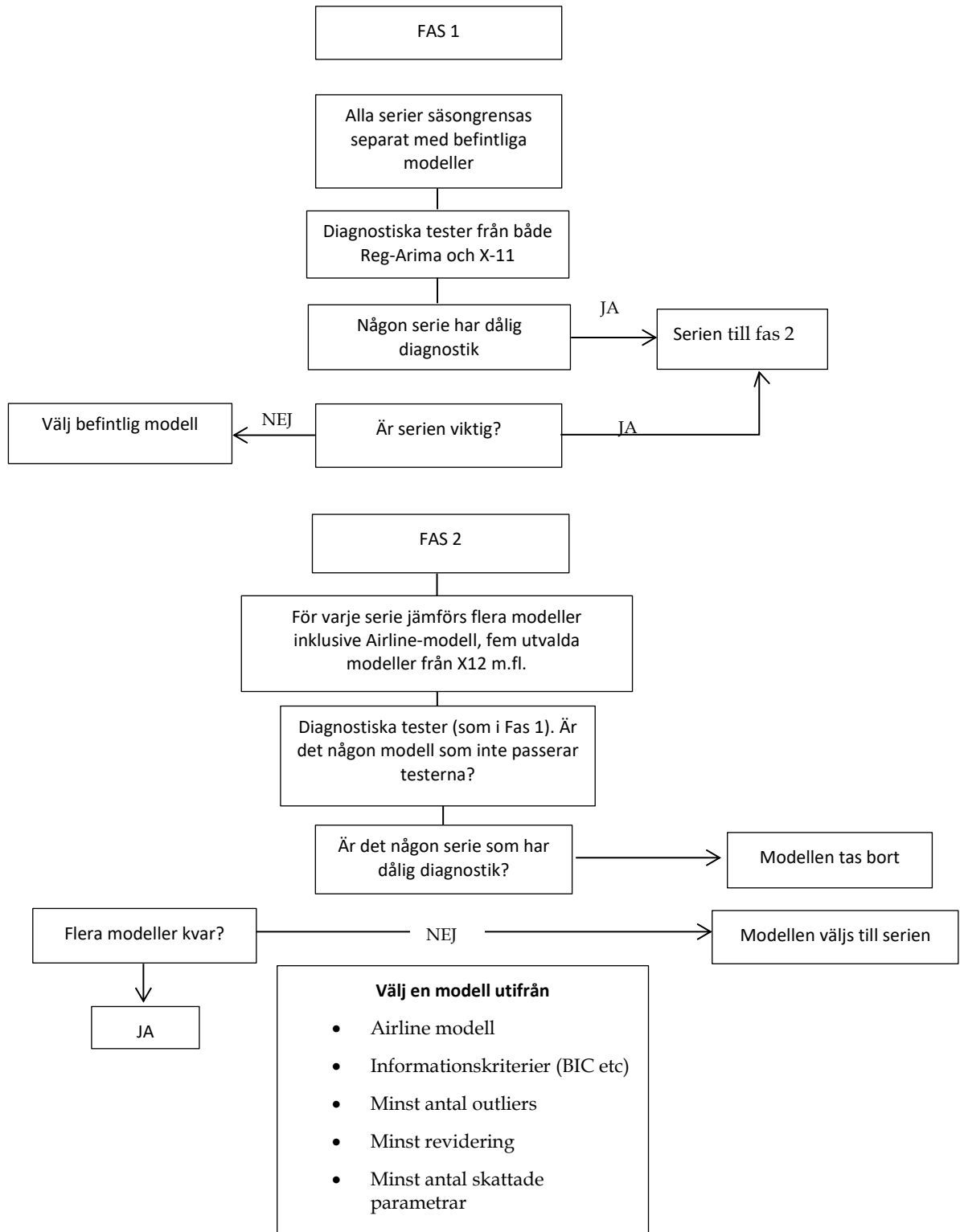


Diagram 3.1 Ett möjligt tillvägagångssätt för modellöversyn.

Innan man börjar testa befintliga och alternativa modeller enligt diagrammet kan det vara lämpligt att höra sig för med ämnessidan (produktansvarig) om aktuella erfarenheter. Ta vara på eventuell information från användarkontakter m.m. huruvida det har kommit fram några problem i användningen av de säsongrensade serierna i fråga.

Fas 1 i *diagram 3.1* innebär att en konservativ ansats föredras. Med andra ord kommer den befintliga modellen att användas även i fortsättningen om den fortfarande fungerar bra när nya observationen läggs till. På så vis behålls stabilitet samtidigt som revideringar reduceras.

Modeller som inte längre beskriver data på ett tillfredställande sätt ska bytas ut. Modellöversynen bygger på ett antal diagnostiska tester, testerna är till stor del inbyggda i SAS-proceduren X12 och gäller för både RegARIMA- och X11- delen. Kriterier för modellöversyn beskrivs mer utförligt i avsnitt 2.6.1 och avsnitt 4.4.

Ifall en befintlig modell inte klarar de diagnostiska testerna – samtidigt som det finns en annan modell som klarar testerna och som har minst lika bra andra egenskaper – borde den befintliga modellen bytas ut mot den alternativa modellen. Den här typen av bedömningar följer inte någon exakt mall utan förutsätter både erfarenhet och goda metodologiska kunskaper.

Resultat från diagnostiska tester och olika utvärderingskriterier i PROC X12 är inte sammanställda på ett ställe vilket försvårar analysen. En stor del av diagnostiken måste specificeras separat i programkoden för att resultaten ska visas.

3.3.3 Modellöversyn i praktiken

Nedan redovisas arbetsgång för modellöversyn avseende det uppsatta provsystemet för IPI med hjälp av TRAMO/SEATS för Windows (TSW). Anledningen till att TSW används i exemplet är att TSW sammanställer utvärderingskriterierna i en tabell automatiskt, något liknande är inte möjligt i PROC X12.

I ett första steg genomförs säsongrensning med gällande modellinställningar, utvärderingskriterierna från TRAMO-delen sammanställs sedan, dessa redovisas i tabell 3.1 nedan. För varje serie redovisas transformation (Lam, 0 som logaritmerad och 1 inte), konstanttermen (Mean), ARIMA-modell (P, D, Q, BP, BD, BQ), standard avvikelse (SE), BIC, diagnostik om autokorrelation i residualer (Q-Val), Normalitetstest (N-test, SK, och KUR), tester för kvarvarande säsongvariation i residualer (QS), autokorrelation i kvadrerade residualer (Q2) och teckentest för residualer (RUNS).

Mått/Serie	8	1012	16_1	22	26	33	B	B_C	C	VKON
Lam	0	1	1	1	0	1	1	0	0	0
Mean	1	0	0	1	1	0	0	0	0	0
P	2	0	0	2	0	2	0	0	0	1
D	0	1	1	1	1	1	1	1	1	1
Q	0	1	1	0	1	0	2	1	1	0
BP	0	0	0	0	0	0	1	0	0	0
BD	1	1	1	1	1	1	0	1	1	1
BQ	1	1	1	1	1	1	0	1	1	1
SE(res)	0,09	2,89	5,82	4,53	0,10	5,90	5,29	0,03	0,03	0,06
BIC	-4,70	2,20	3,65	3,15	-4,47	3,76	3,50	-7,00	-6,98	-5,53
Q-val	18,6	36,7	39,8	45,1	12,2	29,4	24,3	23,7	25,1	47,5
N-test	0,7	1,1	1,9	2,8	0,8	5,9	3,4	1,4	1,3	0,6
SK(t)	0,6	0,8	-1,0	-1,4	-0,3	0,9	-1,8	-1,1	-1,1	-0,7
KUR(t)	0,6	-0,6	1,0	0,9	-0,9	2,2	-0,2	-0,5	-0,3	0,4
QS	0,1	3,8	0,0	1,8	0,0	0,0	0,0	0,0	0,0	0,0
Q2	33,9	28,1	35,4	22,7	16,5	104,4	34,9	31,8	31,6	20,1
RUNS	0,3	-0,5	-0,7	-0,2	0,2	-1,8	-0,6	0,5	1,5	0,8

Tabell 3.1 Utvärderingskriterier med nuvarande modellinställning och serier i IPI provsystem

För att tolka värdena i tabellen krävs kunskap om de teoretiska fördelningarna för de ingående statistikerna och kritiska gränser. Då TSW används kan outputn från programmet tolkas med ett Excel-verktyg från Spaniens Centralbank (se dokument och nerladdning [här](#)). Verktuget flaggar modellerna för SNI_22 och SNI_VKON eftersom värdet på Q-val överstiger den kritiska gränsen, ett tecken på autokorrelation i residualerna. Även modellen för SNI_33 flaggas, värdet för Q2 överstiger den kritiska gränsen och samtidigt klarar inte modellen N-testet. För övriga modeller har det fungerat bra.

De tre problematiska serierna utreds vidare. I första hand jämförs "gamla" modellen med Airline-modellen och därefter med de modeller som kan väljas automatiskt (från X12 eller TSW).

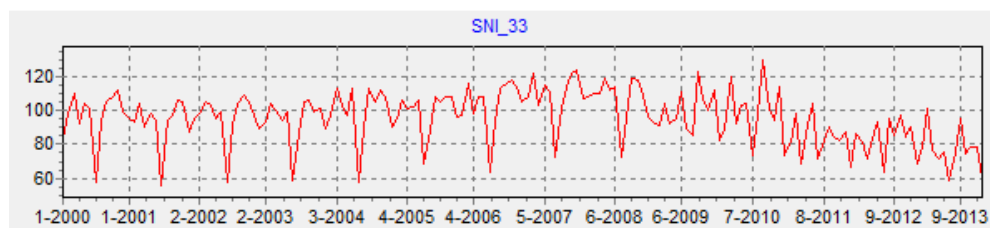
Det fortsatta arbetet exemplifieras med serien SNI_22, i tabell 3.2 nedan visas diagnostika för fem olika modeller. Notera att BIC och SE inte kan användas för att jämföra additiv och multiplikativ modell. Alla modeller, förutom modell tre, har problem med något av de diagnostiska måtten. Men modell tre innehåller fyra outliers, medan den nuvarande inställningen inte har någon outlier. Om modell tre skulle användas så skulle det innebära en stor revidering av den säsongrensade serien, något som skulle kunna ses som problematiskt från ämnessynpunkt. Efter att man vägt samman för- och

nackdelar med att byta modell väljs den modell som använts tidigare för att säsongrensa serien.

Modell/Mått.	Lam	Mean	P	D	Q	BP	BD	BQ	SE(res)	BIC	Q-val	N-test
1	1	1	2	1	0	0	1	1	4,53	3,15	45,09	2,79
2	0	0	0	1	1	0	1	1	0,04	-6,27	38,87	12,80
3	0	0	2	1	1	0	1	1	0,04	-6,35	39,78	1,54
4	1	0	0	1	1	0	1	1	4,50	3,09	48,96	2,44
5	1	0	3	1	1	0	1	1	4,22	3,09	53,78	0,45

Tabell 3.2 Jämförelse för serien SNI_22.

Det visar sig även vara svårt att hitta en alternativ modell till serien SNI_33. Ingen modell lyckas hantera problemet med autokorrelation i de kvadrerade residualerna, mått Q2. En anledning kan vara att säsongmönstret förändrats under perioden (se figur 3.3 nedan). Ett sätt att möta det kan vara att dela serien i två perioder och modellera de två perioderna separat, något som inte görs här. Eftersom problemet med autokorrelation inte kan lösas så behålls den "gamla" modellen.



Figur 3.3 Serien SNI_33.

För serien SNI_VKON, visar det sig att Airline-modellen ger bättre diagnostik än nuvarande modell, Q-val befinner sig inom den kritiska gränsen. Och när outlierstrukturen studeras framgår att ingen outlier förekommer i någon av modellerna.

Serie/Mått	Lam	Mean	P	D	Q	BP	BD	BQ	SE(res)	BIC	Q-val	N-test	Q2
"VKON"	0	0	1	1	0	0	1	1	0.06	-5.53	47.50	0.61	20.12
"VKON"	0	0	0	1	1	0	1	1	0.06	-5.60	31.44	0.13	29.26

Tabell 3.3 Jämförelse för serien SNI_VKON

3.3.4 Efter modellöversynen

En slutlig utvärdering ska normalt göras efter att en modellinställning för varje serie i systemet har valts. Detta innebär att en testkörning av

säsongrensningen görs med nya modellinställningar. Utvärderingen görs enligt principerna i avsnitt 2.6.

3.4 Löpande stöd för säsongrensningen

Vid löpande produktion under året, när det inte ska göras någon modellöversyn, ligger ansvaret att köra säsongrensningsprogram och att granska resultat hos ämnespersonal. Vid oväntade fel eller misstänkta sådana ska ämnesansvariga kontakta ansvariga säsongrensningsspecialister som ska undersöka frågan och eventuellt åtgärda problemet.

Säsongrensningsspecialister får vid behov bevaka utförandet av säsongrensningen. Man behöver kontrollera om det finns felmeddelanden i SAS-loggen, om det finns stora svängningar i säsongrensade resultat. Här är det viktigt att studera utvecklingstal, uppstår oförklarliga utvecklingar åt olika håll, förekommer nivåändringar i trendsattningar och andra liknande avvikelser. En översiktlig diagnostik, förslagsvis i form av diagram är att rekommendera. Sådana övervakningsåtgärder kan delvis överlåtas till ämnesansvariga genom t.ex. grafiska verktyg för granskning, i eller utanför säsongrensningssystemet.

Säsongrensningsspecialister kan vid förfrågan bidra i arbetet med pressmeddelande eller i situationer där de säsongrensade resultaten behöver ytterligare förklaringar. Det ligger också i säsongrensningsspecialisters ansvar att svara på mer komplicerade frågor från användare när det gäller säsongrensning.

4 Bilagor

4.1 Återanvändning av Vägledningens provsystem

Ett alternativ då ett nytt system för säsongrensning sätts upp, är att använda det provsystem som beskrivits i vägledningen. Arbetsgången skulle då vara att kopiera hela mappstrukturen till en produktionsmapp för produkten och sedan anpassa SAS-programmen till provsystemet. Detta kan göras genom att t.ex. ändra sökvägar till indata, ändra parameterfil, ändra format av parameterfil, ändra output och göra ytterligare nödvändiga anpassningar.

Samtidigt är det värt att notera provsystemet speglar egenskaper hos IPI-systemet, vilket har vissa särskilda inslag. Ifall att man vill kopiera provsystemet till andra produkter måste hänsyn tas till produkternas egna egenskaper och kravställningar. Nedan sammanfattas några av sådana inslag i korta drag.

- Låsning av outliers: IPI-systemet tillämpar låsning av outliers. Dvs., då en outlier identifierats i någon tidpunkt i modellöversynen – och om outlieren bedöms vara signifikant – kommer tidpunkten alltid att behandlas som outlier av samma typ i kommande säsongrensningssomgångar. Den här typen av behandlingen fortsätter på samma sätt till dess att en ny modellöversyn genomförs. Då görs även en ny bedömning av läget kring outlierseffekter. Syftet med låsning av outlier är att undvika att outliers *hoppa in och ut* i olika produktionsomgångar vilket borde leda till minskade revideringar. (se exempel i avsnitt 2.2.2)
- Parameterfilens format: Parameterfilen i provsystemet är från en tidigare version av produktionssystemet som använder programmet TRAMO/SEATS och innehåller vissa benämningar som inte har motsvarighet i PROC X12. Det går att ändra vid behov och sen justera motsvarande egenskaper i SAS-programmen.
- Kalendervariabel: I IPI- och provsystemet används den standardiserade kalendervariabeln, se avsnitt 2.3 och avsnitt 4.2.
- Inbyggd felsökning: I SAS-programmet finns automatiska felsökningar inbyggda. Dels kommer ett fönster att "poppa upp" och förklara den bakomliggande orsaken till ett eventuellt fel; samtidigt sparas en *loggen* för varje produktionsomgång. Om det inte finns behov för komplicerad felsökning går det att ta bort vissa kontroller.
- Output som sparas: I provsystemet för IPI har endast ett begränsat antal output tabeller sparats. Flera tabeller genereras av PROC X12, dessa kan sparas vid behov.
- Skillnad mellan provsystemet och det riktiga IPI-systemet: Provsystemet har enbart tio serier, vissa sekretessbelagda serier

hade tagits bort. I det riktiga IPI-systemet hämtas indata från servern P:\Prod\ES vid varje produktionsomgång. Säsongrensade data i det verkliga IPI-systemet måste göras om till textfiler som sedan laddas upp till Statistikdatabasen. Efterarbete med sekretessbelagda serier görs i det riktiga IPI-systemet.

4.2 Modell för säsongrensning och kalenderkorrigering

4.2.1 RegArima modell

I en något förenklad form kan en regressionsmodell med Arima errors - RegArima (U.S.Census Bureau, 2013) skrivas som

$$Y_t^F = \beta_0 + \sum_{l=1}^L \beta_l KF_{l,t} + \sum_{i=1}^k \omega_i \tau_i(B) D_{i,t} + Z_t, \quad (4.1)$$

där Y_t^F är den faktiska serien som beskriver input till den första delen av säsongrensningsproceduren, och Z_t är serien av residualtermer som antas följa en ARIMA modell av följande slag:

$$\phi(B) \Phi(B^s)(1-B)^d(1-B^s)^D Z_t = \theta(B)\Theta(B^s)\varepsilon_t, \quad (4.1.a)$$

där B är lag- operatoren så att $B^s(X_t) = X_{t-s}$, polynomen $\phi(B) = (1 - \varphi_1 B - \dots - \varphi_p B^p)$ är autoregressiv (AR) lag-operatör för ett bestämt p , $(B) = (1 - \vartheta B - \dots - \vartheta_q B^q)$ är moving average (MA) lag-operatör för ett bestämt q , $\Phi(B^s) = (1 - \Phi_1 B^s - \dots - \Phi_p B^{ps})$ är säsong AR-operatören och på motsvarande sätt är $\Theta(B^s) = (1 - \Theta_1 B^s - \dots - \Theta_q B^{qs})$, säsong MA-operatören för $s \in (4,12)$, d och D är differensoperatörer och ε är serie av slumpstermer som följer vitt brus- processen (oberoende lika fördelade observationer med medelvärde noll och konstant varians). Z_t är output från regressionsdelen (det första delsteget)- och input till den andra (dekomponerings-) delen av säsongrensningsproceduren. Z_t kommer att dekomponeras i komponenterna: säsongeffekt, trend-cykel effekt respektive irreguljär effekt:

$$Z_t = S_t + TC_t + I_t, \quad (\text{om additiv modell}) \quad (4.2)$$

$$Z_t = S_t \times TC_t \times I_t, \quad (\text{om multiplikativ modell}). \quad (4.3)$$

Säsongrensningsverktyg har en inbyggd algoritm som testat vilken typ av modell, additiv eller multiplikativ, som bäst passar till data. Modellen innefattar en eller flera förklaringsvariabler ($l=1, \dots, L$) i form av kalenderfaktorer $KF_{l,t}$, samt ett okänt antal outliers D_i som i princip är dummy-variabler. Polynom $\tau_i(B)$ reflekterar vilken typ av outlier det handlar om.

Första delen av proceduren (RegArima) beskriver anpassning av en regressionsmodell där deterministiska effekter såsom kalenderfaktorer och outliers (extremvärden), estimeras.

Andra delen (dekomponering) använder sig av den lineariserade serien Z_t som beskriver den faktiska serien som fri från störningar, såsom outliers och kalendereffekter. Metoden för dekomponering grundas på X-11 proceduren

som har utvecklats av Statistics Canada. Detaljer kring proceduren kan hittas i SAS/ETS User's Guide (SAS Institute Inc, 2012). På SCB används normalt PROC X12:s defaultfilter som är baserad på Statistics Canadas *moving seasonality ratio* (MSR).

4.2.2 Kommentarer om antal outliers

Gränsen för "alldeles för många outliers" beror på antal observationer och andra verklighetsrelaterade krav (t.ex. naturliga ändringar i tidserier p.g.a. vissa fenomen som går att identifiera i verkligheten). Ett exempel är när variabler byter definition i samband med SNI-omläggningar vilket vanligtvis görs efter en specifik tidpunkt – eller när starka konjunktursvängningar påverkar tidseriens nivå (plötsliga- och mer eller mindre kvarstående förändringar i arbetslöshet, sysselsättning, produktion och andra viktiga makroekonomiska variabler). I vissa programvaror (t.ex. i Tramo-Seats, UCM osv) förekommer tumregler gällande antal outliers som kan ge upphov till problem med modellenpassning. Ibland rekommenderas gränsen vid 5 eller fler outliers – eller alternativt 5% outliers av totala antalet av observationer – men de här reglerna har inte någon vetenskaplig grund utan baseras på empiriska utvärderingar.

4.3 Kalenderkorrigeringsprinciper

SCB:s standardprocedur för kalenderkorrigering innebär att en kalenderfaktorn definieras i form av kvot eller differens av följande slag:

- Om multiplikativ modell för säsongrensning

$$KF_t = \ln \left(\frac{N_{t,m}}{N_m^*} \right); \quad (4.4)$$

- Om additiv modell för säsongrensning

$$KF_t = N_{t,m} - N_m^* \quad (4.5)$$

där $N_{t,m}$ är antal arbetsdagar år t och kvartal/månad m medan N_m^* är ett genomsnittligt antal arbetsdagar för kvartal/månad m över en lång tidsperiod (multipel av 28). $N_{t,m}$ beräknas enligt svensk kalender där röda dagar och veckoslutshelger exkluderas.

Ibland förekommer en förenklad version av det långsiktiga genomsnittet N^* istället för N_m^* . N^* är då ett konstantvärde som beskriver det långsiktiga medelvärdet för antal arbetsdagar per kvartal/månad, oavsett vilket kvartal/månad det handlar om.

Tabellen 4.1 visar ett exempel för beräknade genomsnittliga antal arbetsdagar per kvartal respektive månad över perioden jan1900 – dec 2095, (196 år). Empiriska tester har visat att förändringar i det långsiktiga genomsnittet stabiliseras på en andra decimalnivå redan för perioder längre än ca 100 år. Notera att värdena i tabellen är baserade på en grundläggande

uppsättning av helgdagar. Eventuella förändringar i uppsättningen skulle sannolikt orsaka förändringen i de genomsnittliga värdena i tabellen. Exempelvis i systemet för nationalräkenskaper har ett manuellt beräknat värde, 62,6875, använts för N^* .

Kvartal	N_m^*	N^*
m=1	62,58	62,76
m=2	59,88	62,76
m=3	65,71	62,76
m=4	62,86	62,76

Månad	N_m^*	N^*
m=1	20,72	20,92
m=2	20,17	20,92
m=3	21,69	20,92
m=4	19,88	20,92
m=5	20,17	20,92
m=6	19,83	20,92
m=7	22,14	20,92
m=8	22,14	20,92
m=9	21,43	20,92
m=10	22,14	20,92
m=11	21,43	20,92
m=12	19,29	20,92

Tabell 4.1. Genomsnittliga antal arbetsdagar per kvartal och månad enligt svensk kalender för perioden 1900-2095.

Beräkning av kalenderkorrigerade serier görs enligt de principer som beskrivs nedan.

- Om *multiplikativ* modell för slutlig dekomponering av tidserien används gäller följande

$$KK_t = \left(\frac{Y_t^F}{KF_t^*} \right), \quad (4.6)$$

där KK är kalenderkorrigerad serie, Y^F är den faktiska originalserien och KF^* är skattningen av kalenderfaktorer ur säsongreningsproceduren.

Säsongrensningsverktyg genererar automatiskt variabeln KF^* enligt beräkningsprincipen

$$KF_t^* = \exp\{\hat{\beta}_1 \times KF_t\} \quad (4.7)$$

- Om en additiv modell används, gäller följande

$$KK_t = Y_t^F - KF_t^* \quad (4.8)$$

I det här fallet genererar säsongrensningsverktyget variabeln KF^* enligt

$$KF_t^* = \hat{\beta}_1 \times KF_t \quad (4.9)$$

4.3.1 PROC X12 anpassning

Notera att PROC X12 inte genererar den kalenderkorrigerade serien KK_t per automatik utan enbart skattningar för kalenderfaktorer KF_t^* (tabellen D18 i PROC X12 output).

I praktiken innebär detta att värden för de slutliga kalenderkorrigerade serierna beräknas enligt (4.6) respektive (4.8) på följande sätt:

- För en multiplikativ modell gäller

$$KK_t = \left(\frac{Y_t^F}{D18} \right); \quad (4.10)$$

- Alternativt, för additiv modell gäller

$$KK_t = Y_t^F - D18 \quad (4.11)$$

4.4 Andra kalendervariabler: påskeffekter

4.4.1 Inbyggda påskeffekter

Påskeffekter och eventuella handelsdageffekter kan implementeras om det finns särskilda skäl att göra detta. Ett exempel är månadsstatistik från detaljhandelns försäljning som tydligt påverkas av påskeffekt och/eller sammansättning av veckodagar.

Ämnespersonal bör rådfrågas om det finns tveksamhet kring införandet av påskeffekter i säsongrensningen.

Det enklaste sättet att hantera påskeffekten inom ramen för X-12-Arima är via den inbyggda möjligheten att hantera påsken på ett speciellt sätt enligt principer som U.S. Census Bureau definierat. Då påsken inträffar vid samma tidpunkt som i USA kan proceduren användas utan särskilda anpassningar till svensk kalender.

Påskeffekten representeras av kalenderfaktorn

$$P_t(w, PT) = \frac{m_t}{w} - \mu_{w,t} \quad (4.12)$$

där m_t är antal dagar före Påsken (PT är datum för Påskinträffande) som hamnar i månad t . Notera att w bestäms i förväg genom att restriktionen $1 \leq w \leq 25$ uppfylls.

Variabeln P är alltid lika med noll för alla månader förutom för februari, mars och april. P är skild från noll i februari endast om att $w > 22$. I praktiken väljs $w < 22$ så att februarieffekten inte tas med (våldigt osannolikt att det finns någon februarieffekt för påsken för månadsdata då påsken kan inträffa mellan 22 mars och 25 april). $\mu_{w,t}$ är långsiktigt medelvärde av första delen i (4.12), $\frac{m_t}{w}$, som motsvarar de första 400 åren enligt Gregoriansk kalender, 1538-1982. De här medelvärdena är inbyggda i proceduren och varierar beroende på vilket värde w har (se Findley m.fl., 2005).

Bedömningen av w baseras i huvudsak på praktisk erfarenhet. Då det handlar om skattning inom en regressionsmodell blir valet av värdet inte särskilt betydande om det inte överstiger 22, vilket skulle innebära att påskeffekten blir synlig även i februari. Vanligtvis väljs w mellan 7 och 14 då det finns en uppfattning att påskeffekten blir mest synlig mellan en till två veckor innan inträffandet.

Kalenderfaktorn (4.12) förutsätter att påskeffekten ger upphov till ändring i ekonomisk aktivitet ca w - dagar innan påsk och den här ändringen behåller samma nivå t.o.m. dagen före påsk. Termen $\mu_{w,t}$ säkerställer att säsongeffekten separeras från kalendereffekten.

4.4.2 Exempel

T.ex. om Påsken (PT) inträffar 4 april och $w=8$ blir

$P_{mars}(w = 8, PT = 4apr) = \frac{5}{8} - 0,3534 = 0,2716$, (fem av åtta dagar som faller i mars);

$P_{apr}(w = 8, PT = 4apr) = \frac{3}{8} - 0,6466 = -0,2716$, (tre av åtta dagar som faller i april).

Låt oss anta att multiplikativ modell gäller, vilket innebär att log-transformation implementeras, samtidigt som parameterskattningen för påskeffekt blir 0,0252. I så fall blir den skattade kalendereffekten för mars respektive april

$$KF_{t=mars}^* = \exp\{0,00252 \times 0,2716\} \approx 1,0007,$$

$$KF_{t=apr}^* = \exp\{0,00252 \times -0,2716\} \approx 0,999.$$

Andra månader har kalendereffekten lika med 1 ($\exp(0) = 1$) vilket innebär att Påskeffekten inte har någon påverkan då. Om

$$Y_{t=mars}^F = 75,325 \text{ och } Y_{t=april}^F = 77,925,$$

blir kalenderkorrigerade värden för motsvarande månader beräknade på följande sätt

$$KK_{t=mars} = \frac{75,325}{1,0007} = 75,27,$$

$$KK_{t=apr} = \frac{77,925}{0,999} = 78,003.$$

4.4.3 Egen påskvariabel

I Findley m.fl., (2005) diskuteras andra typer av kalenderfaktorer för de avvikande situationerna när man t.ex. behöver separera effekter på ett antal dagar före påskhelgen och helgdagar. De här kalenderfaktorerna är dock inte implementerade i X-12-Arima vilket innebär att de måste skapas som externa förklaringsvariabler.

En sådan variabel måste innehålla värden som motsvarar längden av tidserien som säsongrensas inklusive längden för prognosvärden. Rekommenderat antal observationer är $T+3$ år, där T är totala antalet tillgängliga observationer i tidserien.

Den här variabeln borde definieras som en dummy variabel om man vill att effekten ska skattas på det sättet som U.S. Census Bureau använder som standard i X-12-Arima och PROC X12.

I PROC X12 måste variabeln specificeras som en regressionsvariabel genom att ange kommando:

```
Regression Uservar = (variabelnamn) / Usertype =  
Easter;
```

Det går att utesluta delen *Usertype= Easter* vilket skulle innebära att programmet behandlar variabeln som en vanlig regressionsvariabel och inte som påskeffekt enligt metoden som U.S. Census Bureau använder.

Alternativet till ovanstående kommando är kommando *Event* som har exakt samma betydelse som kommando *Regression* i det här sammanhanget:

```
Event (variabelnamn) / Usertype = Easter;
```

4.4.4 Andra kalendereffekter: skottår och månads-/kvartalslängd

Ibland uppstår behov att tillgodoräkna effekten av skottår som vanligtvis görs via de inbyggda specifikationen:

```
Regression Predefined = LpYear;
```

Notera att skottåreffekten kan vara i konflikt med andra relaterade kalenderfaktorer vilket medför att det inte går att använda specifikationen ovan tillsammans med några andra specifikationer (se avsnittet *The X12 Procedure* i SAS ETS 12.1 User's Guide/Procedure Reference/).

4.5 Exempel SAS programkod

I nedanstående exempel (figur 4.1) visas specifikation för serier SNI_B_C, SNI_C och SNI_B med automatisk modellering, automatisk outlierhantering och kalenderfaktor *ln_kalf* som låst förklaringsvariabel. Även typ av transformation sköts via programmets inbyggda mekanism som specificeras med kommandot

```
Transform function = Auto.
```

Hade modellen valts manuellt genom att ange egna preferenser hade val av multiplikativ - respektive additiv dekomponering specificerats med följande uttryck:

```
Transform function = Log;...X11 mode = Mult;
```

respektive

```
Transform function = None;...X11 mode = Add;
```

Förklaringsvariabeln är regressionsvariabel i RegArima-delen som specificeras med kommando:

```
REGRESSION USERVAR= Ln_Kalf/ USERTYPE= LOM ;
```

där `Usertype = LOM` försäkrar att kalenderfaktorn kommer att behandlas på samma sätt som en inbyggd kalendervariabel.

Programmet kommer att hantera outliers på ett automatiskt sätt för hela tidserien (från januari 2000 och framöver), vilket anges genom specifikationen

```
OUTLIER SPAN=( 'Jan2000', )
          TYPE=( AO TC LS ) ;
```

Avaktivering av kommando *Outlier* skulle leda till att programmet inte tar hänsyn till outlierseffekter. Med andra ord, varken skattning eller korrigering för dessa effekter skulle bli genomförd. I de flesta fall skulle detta innebära försämrade kvalitet och i sällsynta fall skulle procedurernas algoritmer inte konvergera vilket skulle leda till att tidserier inte blir säsongrensade.

4.6 Utvärdering av säsongrensningmodeller

4.6.1 Jämförelse mellan preliminärt valda modeller

- Jämför 5 *bästa* modellalternativ per serie: Finns signifikanta skillnader i likelihoodbaserade kriterier (informationskriterier)?
 - tumregel som ibland kan användas: om det finns skäl att välja modell 2 framför modell 1(bäst), kontrollera om $\Delta = AIC_{alt} - AIC_{opt} < 2$. Om $\Delta \leq 2$ får alternativ modell väljas. Om $3 < \Delta < 8$ kan den alternativa modellen övervägas om det finns starka skäl, annars normalt inte. Ifall $\Delta > 10$ är den alternativa modellen inte tillräckligt bra.
- Jämför modeller med- och utan outlierhantering/kalenderkorrigering: förändras då automatiskt valda modellinställningar (tyder någonting på instabilitet)?
 - Jämför ett par *bästa* modeller (per serie) som innehåller outlierhantering (och/eller kalenderkorrigering ifall det är aktuellt):
 - Är kriteriemåttens värden nära varandra? Observera: likelihoodbaserade mått kan användas enbart för jämförelse mellan modeller med samma

outlierpositioner och samma typ av transformation (*logaritmisk alt. ingen*).

- Är parameterskattningarna signifikanta? t-tester m.m. Vilken modell har minst antal parametrar? (färre parametrar är att föredra, allt annat lika – eng. *parsimonious model*).
 - Vilken modell orsakar minst volatila förändringar i trend&säsongrensad serie? (*jämnhetsfråga, dock inte alltid- då jämna serier kan vara revideringsbenägna*).
 - Vilken modell har minst antal identifierade outliers (viktigt)
- Jämför *average forecast error* (om stora avvikelser finns mellan två eller flera modeller).
 - Ytterligare kriterier, t.ex. revideringsbenägenhet (kör automatisk modellering på en kortare period och jämför modellstabilitet med hänsyn till hela perioden).

4.6.2 Säsongdekomponeringen

- Är säsongvariationen signifikant? (kontrollera resultat från test för identifiable seasonality som beskrivs i ETS guide).
- Diagnostiska mått M1-M11 (M7 viktigast). Mått som överstiger 1 tyder på eventuella problem med skattning i förhållande till kriteriet som måttet relaterar till.
- Finns kvarvarande säsongvariation i säsongrensade serier & irreguljär komponent? Undersök genom att säsongrensa den serie som precis säsongrensats/irreguljär komponent. Om PROC X12 indikerar att det finns kvarvarande säsongmönster är det ett tecken på att modellen kan behöva ses över.

4.6.3 Andra kriterier

Jämför fler modeller per serie: konsistens, logiska utvecklingar och ytterligare aspekter som kan vara viktiga. Exempel på viktiga kriterier:

- Revideringsstorlek: små revideringar föredras. Gör modellanpassningar med samma modell på delperioder och undersök om outlierstruktur förändras. Gör automatiska modellanpassningar på delperioder och undersök modellstabilitet (om programmet identifierar samma ARIMA-modell, om outliereffekter identifieras på samma positioner, kalendereffekter mm).
- Modellers förmåga att fånga "sanna" outliers kan utvärderas (t.ex. outliers som modellen borde hitta i serien p.g.a. ekonomiska orsaker såsom konjunkturedgångar, naturliga förändringar i serien p.g.a. t.ex. SNI-omläggningar eller andra särskilda skäl). Om serierna uppfattas som särskilt brusiga kan det vara lämpligt

att avväga användning av en och samma RegArima-modell för samtliga serier i systemet (företrädesvis Airline-modell). Brus i tidserien kan tolkas som oförklarlig variation. En alltför brusig tidserie innehåller så mycket brus i förhållande till signalen att det inte går att få tillräckligt stabila skattningar av parametrarna i modellerna. En brusig tidserie brukar ha stark revideringsbenägenhet, stor variation kring medelnivån på långsikt (trend), varierande säsongvariation, många outliers osv. Signalen i säsongrensningssammanhanget kan i någorlunda förenklad form förklaras som: trend + förklarliga linjära och icke linjära effekter (outliers, kalendereffekter) = säsongrensad serie. Bruset får anses innefatta allt som inte kan förklaras med vår linjära modell, t.ex. oregelbunden säsongvariation, oidentifierade outliers eller interaktioner mellan trendcykel och säsong.

- Den ovan nämnda principen om en RegArima-modell för hela systemet av tidserier kan också vara aktuell om stark *moving seasonality* förekommer.
- Eventuellt kan statistikredovisningen då behöva kompletteras med varningar till statistik användare, om risker i användningen.
- När det är motiverat kan låsning av outlierpositioner användas för att inte bryta interna konsistenser mellan relaterade serier.
- Jämför utvecklingar i kalenderkorrigerade-, trend-cykel och säsongrensade serier med olika modellalternativ i förhållande till ämnesspecifika kvalitetskrav.

```
TITLE 'X12-ARIMA FÖR IPI med Kalenderfaktor & outlierkorrigering';

ODS LISTING ;
ODS OUTPUT F3= Modl_01_KVAL;
ODS OUTPUT D8A= Modl_01_Diagnostika;
ODS OUTPUT ForecastCL= Modl_01_Prognos;
ODS OUTPUT MLESUMMARY= Modl_01_MLIKH;
ODS OUTPUT G = Modl_01_SPktr;
ODS OUTPUT ErrorACF= Modl_01_ACF_resid;
ODS OUTPUT ErrorPACF= Modl_01_PACF_resid;
ODS OUTPUT BEST5MODEL = Modl_01_Best5;
ODS OUTPUT AUTOCHOICEMODEL = MODL_AUTOCHOICEMOD;
ODS OUTPUT AutomaticModelChoice= Modl_01_jamfor;
ODS OUTPUT UNITROOTTESTMODEL = Modl_01_URTESTMDL;
ODS OUTPUT FINALMODELCHOICE = Modl_01_FINALMOD;
ODS OUTPUT REGPARAMETERESTIMATES = Modl_01_REGEST;
ODS OUTPUT ARMAParameterEstimates= Modl_01_ARmaParEst;
ODS OUTPUT ARMAIterationSummary= Modl_01_ARMA_iter;
ODS OUTPUT RESIDUALHISTOGRAM = Modl_01_REGARMRESID;
ODS OUTPUT OUTLIERDETECTION = Modl_01_OUTLIER_CV;
ODS OUTPUT NormalityStatistics = Modl_01_NORMSTAT;
ODS OUTPUT RESIDUALSTATISTICS= Modl_01_RESSTAT;
ODS OUTPUT AVGFCSTERR = Modl_01_AVGFCSTERR ;
ODS OUTPUT OUTLIERDETECTION= Modl_01_CVOUTL ;
ODS OUTPUT ForecastCL = Modl_01_ForcstCL;

PROC X12 AUXDATA= WORK.Kalfaktor
DATA= WORK.IPIserier_201402
SPAN=('Jan2000', )
DATE= DATUM
INTERVAL= Month
MDLINFOOUT= WORK.Modl_01_IPI
OUTSTAT= Modl_01_OutStat;
VAR SNI_B_C SNI_C SNI_B ;
TRANSFORM function= Auto;
CHECK PRINT=(ACF PACF ACFSQUARED
NORM RESIDUALOUTLIER
RESIDUALSTATISTICS
SPECRESIDUAL);
REGRESSION USERVAR= Ln_Kalf/ USERTYPE= LOM;
AutoMdl
print = AutoChoiceMdl
Best5Model
UnitRootTest UnitRootTestMdl ;

ESTIMATE;
FORECAST;
OUTLIER SPAN=('Jan2000', )
TYPE=( AO TC LS ) ;
X11;
OUTPUT OUT= WORK.IPI_Modl_01_OUTSER
A1 A8 B1 D10 D11
D12 D13 D18 E6
;

RUN;
QUIT;
ODS LISTING CLOSE;
```

Figur 4.1 SAS program för säsongrensning för IPI.

5 Litteraturförteckning

- Banco de España. (maj 2011). *Banco de España*. (Bank of Spain)
<http://www.bde.es/servicio/software/econome.htm>. Läst den 29 januari 2015.
- Eurostat. (2009). ESS guidelines on seasonal adjustment.
http://epp.eurostat.ec.europa.eu/portal/page/portal/product_details/publication?p_product_code=KS-RA-09-006. Läst den 23 februari 2011.
- Eurostat. (2011). Guidelines for compiling the monthly Index of Production in Construction. European Union.
- G. Caporello, A. M. (2004). *Program TSW: Revised Reference manual*. Banco de España:
<http://www.bde.es/servicio/software/programase.htm>. Läst den 08 april 2011.
- D. Findley, K.C. Wills, B. Monsell. (ASA proceedings, November 2005). *Issues in Estimating Easter Regressors Using RegARIMA with X-12-ARIMA*. Hämtat från U.S. Census Bureau:
<https://www.census.gov/srd/www/sapaper/sapaperbytopic>. Läst den 12 februari 2015.
- H. Lindkvist, H. Pettersson, R. Zetterman. (2012). Byggproduktionsindex- underlag till Riksbankensrapport. Stockholm: SCB.
- Maravall, A. (05 2011). *BANCO DE ESPAÑA: Programs*. Hämtat från Brief description of the programs:
http://www.bde.es/f/webbde/SES/servicio/Programas_estadisticos_y_econometricos/Programas/ficheros/summprogs.pdf. Läst 17 dec 2014.
- Mazzi G. L., Calizzani C. (2009). ESS guidelines on seasonal adjustment. Luxembourg: Eurostat.
- SAS Institute. (2010). *SAS/ETS 9.22 Users Guide*. (SAS Institute)
<http://support.sas.com/documentation/922/index.html>. Läst den 10 maj 2011.
- SAS Institute Inc. (2012). *SAS/ETS® 12.1 User's Guide*. Cary, NC, USA.
- SCB. (den 04 06 2003). *Säsongrensning*.
http://www.scb.se/Pages/Standard_35725.aspx. Läst den 23 september 2011.
- SCB. (2008). *Bakgrundsfakta 2008:6: Kvartalsmönster i BNP, slutrapport*. Stockholm: Statistiska centralbyrån.
- SCB. (2009-2010). *Resultatrapport: Summakonsistent säsongrensning etapp 1 och 2*. Stockholm: Statistiska centralbyrån.
- SCB. (2010). *Resultatrapport: Förstudie för vidareutvecklad säsongrensning*. Stockholm: Statistiska Centralbyrån.

- SCB. (2010-2011). *Resultatrapport: Principer för outlierhantering i säsongrensning- etapp 1*. Stockholm: Statistiska centralbyrån.
- SCB. (2014). *Resultatrapport: Principer för outlierhantering i säsongrensning- etapp 2*. Stockholm: Statistiska centralbyrån.
- U.S. Census Bureau. (2010). *The X-12-ARIMA Seasonal Adjustment Program*. <https://www.census.gov/srd/www/x12a/>. Läst den 02 december 2014.
- U.S. Census Bureau. (den 28 februari 2011). *U.S. Census Bureau*. Hämtat från X-12-ARIMA Reference Manual: <http://www.census.gov/ts/x12a/v03/x12adocV03.pdf>. Läst den 03 december 2014.
- U.S.Census Bureau. (den 24 Oktober 2013). *Census.gov*. X-12-ARIMA Reference Manual Version 0.3 Hämtat från <https://www.census.gov/ts/x12a/v03/x12adocV03.pdf>. Läst den 18 februari 2015.
- V. Gomez, A. Maravall. (november 1997). *Programs TRAMO and SEATS*. http://www.bde.es/f/webbde/SES/servicio/Programas_estadisticos_y_econometricos/Programas/ficheros/manualdos.pdf. Läst den 17 december 2014