

Traballo Fin de Máster

Corrección de series económicas de Galicia

Alba Rodríguez García

Máster en Técnicas Estatísticas

Curso 2021-2022

Proposta do Traballo Fin de Máster

Título en galego: Corrección de series económicas de Galicia

Título en español: Corrección de series económicas de Galicia

English title: Correction of Galicia's economic series

Modalidad: Modalidade B

Autora: Alba Rodríguez García, Universidade de Santiago de Compostela

Directora: María José Lombardía Cortiña, Universidade da Coruña

Tutora: María Esther López Vizcaíno, Instituto Galego de Estatística (IGE)

Breve resumo do traballo:

Revisión metodolóxica do axuste estacional das series de conjuntura económica de Galicia más relevantes do Banco de series de conjuntura, xunto co axuste de novos modelos empregando o software JDemetra+. Ademais, adaptando as ferramentas de JDemetra+ a linguaxe R obterase un procedemento automático co cal se actualizará a Base de datos.

Dona María José Lombardía Cortiña, categoría 1 da Universidade da Coruña, dona María Esther López Vizcaíno, cargo 1 de Instituto Galego de Estatística (IGE), informan que o Traballo Fin de Máster titulado

Corrección de series económicas de Galicia

foi realizado baixo a súa dirección por dona Alba Rodríguez García para o Máster en Técnicas Estatísticas. Estimando que o traballo está rematado, dan a súa conformidad para a súa presentación e defensa ante un tribunal.

En Santiago de Compostela, a 13 de xuño de 2022.



A directora:
Dona María José Lombardía Cortiña

A titora:
Dona María Esther López Vizcaíno



A autora:
Dona Alba Rodríguez García

Declaración responsable. Para dar cumprimiento á Lei 3/2022, de 24 de febreiro, de convivencia universitaria, referente ao plaxio no Traballo Fin de Máster (Artigo 11, Disposición 2978 del BOE núm. 48 de 2022), **A autora declara** que o Traballo Fin de Máster presentado é un documento orixinal no que se tiveron en conta as seguintes consideracións relativas ao uso de material de apoio desenvolto por otros/as autores/as:

- Todas as fontes usadas para a elaboración deste traballo foron citadas convenientemente (libros, artigos, apuntes de profesorado, páxinas web, programas,...)
- Calquera contido copiado ou traducido textualmente púxose entre comillas, citando a súa procedencia.
- Fíxose constar explicitamente cando un capítulo, sección, demostración,... sexa unha adaptación casi literal dalgúnha fonte existente.

E, acepta que, se se demostrará o contrario, se lle apliquen as medidas disciplinarias que correspondan.

Índice xeral

Resumo	IX
1. Estado da arte	1
2. Metodoloxía	9
2.1. Linealización	11
2.1.1. Estabilidade en media e varianza	12
2.1.2. Eliminación dos efectos deterministas	12
2.1.3. Identificación automática do modelo ARIMA	16
2.2. Estimación das compoñentes principais	18
2.2.1. SEATS	19
2.2.2. X-11	23
2.3. Comparación TRAMO-SEATS e X-13ARIMA-SEATS	26
2.4. Diagnose do modelo	27
3. Resultados	33
3.1. Poboación ocupada en Galicia (EPA)	34
3.2. Índice de producción industrial (IPI)	37
3.3. Transporte marítimo.	42
3.4. Ocupación hoteleira: viaxeiros entrados.	46
3.5. Xeración de enerxía eléctrica.	51
4. Automatización do proceso en R	57
4.1. Calendario de Galicia	58
4.2. Especificacións JDemetra+	59
4.3. Exemplo coa serie do índice de producción industrial de Galicia (IPI15XE)	62
4.4. Series na Base de datos	69
5. Conclusións	73
A. Táboa de parámetros dos modelos finais	75
B. Código R	81
Bibliografía	119

Resumo

Resumo en galego

O Banco de series de conjuntura do Instituto Galego de Estatística (IGE) é unha aplicación disponible na web do IGE que contén un conxunto de series de datos estatísticos de diversas fontes de información, que teñen como obxectivo o seguimento da conjuntura económica galega. Co obxectivo de incrementar os servizos que xa provía o Banco de series de conjuntura, fíxose unha selección daquelas series que se consideraban máis relevantes para seguir a evolución da economía galega e incluíronse as correspondentes series corrixidas estacionalmente. Os obxectivos do axuste estacional son identificar e eliminar as fluctuacións estacionais e os efectos de calendario que poidan enmascarar os movementos a curto e longo prazo nunha serie temporal e impedir unha comprensión clara dos fenómenos subxacentes. Polo tanto, o axuste estacional é un proceso fundamental na interpretación de series temporais para a elaboración de políticas públicas. Os obxectivos deste traballo son facer unha revisión metodolóxica dos métodos empregados para realizar as correccións de estacionalidade e calendario que agora mesmo están disponibles no Banco de series de conjuntura, así como, estudar a creación de novos modelos de corrección das series que quedarán implementados mediante algoritmos automáticos para o posterior axuste e a grabación na Base de datos.

English abstract

The Conjecture series bank of the Galician Institute of Statistics (IGE) is an application available in the web of the IGE that contain a set of series of statistical data from diverse sources of information, whose aim is monitoring the Galician economic conjecture. In order to increase the services already provided by the Conjecture series bank, there is a selection of the most relevant series to follow the evolution of the Galician economy, providing them with the corresponding seasonally adjusted series. The goals of seasonal adjustment are to identify and eliminate seasonal fluctuations and calendar effects that may mask short and long term movements in a time series and prevent a clear understanding of the underlying phenomena. Therefore, seasonal adjustment is a fundamental process in the interpretation of time series for the elaboration of public policies. The objectives of this work are to review the methods used to make the seasonality and calendar corrections that are currently available in the Conjecture series bank, as well as to study the creation of new series correction models that are going to be implemented using automatic algorithms. After that, the models are prepared to be adjusted and recorded in the Database.

Capítulo 1

Estado da arte

O Instituto Galego de Estatística (IGE) é un organismo autónomo da Xunta de Galicia creado no ano 1988 e que se rexé basicamente pola Lei 9/1988 de Estatística de Galicia. Na súa misión de promover o desenvolvemento do sistema estatístico da Comunidade Autónoma debe prestar servizos de recompilación e difusión da documentación estatística dispoñible, analizar as necesidades e a evolución da demanda de estatísticas e asegurar a súa difusión.

O seu principal obxectivo é o desenvolvemento do sistema estatístico de Galicia de modo que sirva como fonte de recopilación e difusión da información estatística dispoñible, desenvolvendo bases de datos de interese público, que respondan ás necesidades e á evolución da demanda estatística. Búscase que todos os cidadáns poidan acceder de forma sinxela a ditos datos, garantizando a oportunidade de consultar información precisa sobre diversos ámbitos. En particular, neste traballo centrámonos na análise das series de conjuntura económica de Galicia.

As series de conjuntura económica son conjuntos de datos medidos no tempo, que abranguen diferentes ámbitos relacionados coa economía, nos que se nos achega información acerca da súa evolución en función das circunstancias que se foron dando ao longo do tempo. As series de conjuntura teñen diversas aplicacións, dende a análise e interpretación dos datos ata predicións futuras.

As series brutas de datos, medidas ao longo do tempo, en períodos superiores a un ano, polo xeral, presentan fenómenos que impiden a súa correcta análise ou dificultan a súa interpretabilidade. Exemplo disto son a presenza de estacionalidade, o efecto calendario e a ciclotendencia. Definamos máis detalladamente cada un destes conceptos.

- A estacionalidade dunha serie temporal asóciase á variación periódica e previsible da mesma con un período inferior ou igual a un ano.
- O efecto calendario denota o impacto dos días traballados, laborais, ano bisesto e outros fenómenos asociados á composición do calendario, nunha serie temporal.
- A ciclotendencia divídise en dúas compoñentes. Por un lado temos a serie de tendencia, que recolle os movementos de longa duración, ou baixa frecuencia, cuxo período é superior aos 8 anos. Por outro lado, a compoñente cíclica recolle as oscilacións cuxa duración se sitúa entre 2 e 8 anos. Dada a dificultade para discriminar entre oscilacións de tendencia e cíclicas, é habitual traballar cunha compoñente mixta de ciclo e tendencia, que denotamos por tendencia.

O *Banco de series de conjuntura* do IGE é unha aplicación dispoñible na web do IGE que contén un conxunto de series de datos estadísticos de diversas fontes de información que axudan a explicar a situación económica actual de Galicia. Estes datos están nun formato que permite que sexan consultados por un público de moi diferente índole, dende cidadáns que busquen información ata profesionais que usen os datos para investigación a partir deles. Co obxectivo de incrementar os servizos que xa provía o Banco de series de conjuntura, fíxose unha selección daquelas series que se consideraban máis relevantes para seguir a evolución da economía galega

e incluíronse as correspondentes series corrixidas de estacionalidade e calendario, así como a serie de tendencia asociada, que nos permiten apreciar a evolución económica a longo prazo (máis de dous anos).

Faremos unha selección das series, que polo tipo de información que recollen ou ámbito no que se enmarcan, son consideradas as máis relevantes para explicar a conxuntura económica de Galicia. As series seleccionadas para facer a corrección de estacionalidade e calendario, e obter a súa serie de ciclotendencia, clasificadas por temas (Instituto Galego de Estatística, 2022), son as seguintes:

- Mundo laboral
 - AMSS: Media mensual de traballadores afiliados en alta laboral na Seguridade Social segundo réximes.
 - PAROREG: Paro rexistrado. Total.
 - EPA11_OCUT: Ocupados da Enquisa de poboación activa.
 - EPA11_PART: Parados da Enquisa de poboación activa.
 - EPA11_INA: Inactivos da Enquisa de poboación activa.
- Industria e enerxía
 - IPI15XE: Índice de producción industrial. Índice xeral. Base 2015.
 - PRD_VEH: Produción de vehículos automóbiles.
 - CPPGASOLT: Consumo de gasolina.
 - CPPGO: Consumo de gasóleo.
 - GEETO: Xeración de enerxía eléctrica. Total.
 - ENRTO: Xeración de enerxía eléctrica non renovable. Total.
 - ERTO: Xeración de enerxía eléctrica renovable. Total.
- Construción
 - VIINI: Vivendas iniciadas. Total.
 - VITER: Vivendas terminadas. Total.
 - EDTVIV: Edificación de nova planta. Número de vivendas. Total.
 - EDTSUP: Edificación de nova planta. Superficie a construír. Total.
- Servizos de mercado. Xeral
 - SERV_CN_2015: Indicadores de actividade no sector servizos. Base 2015. Índice xeral da cifra de negocios.
 - SERV_OCUP_2015: Indicadores de actividade no sector servizos. Base 2015. Índice de ocupación.
- Servizos de mercado. Transporte
 - TAP: Transporte aéreo. Pasaxeiro. Total.
 - TMM: Transporte marítimo. Mercancías cargadas, descargadas e transbordadas. Total.
- Servizos de mercado. Movemento turístico
 - EOHP_04: Ocupación nos establecementos hoteleiros. Pernoitas.
 - EOHVE_04: Ocupación nos establecementos hoteleiros. Viaxeiros entrados.
 - EOGRH_04: Ocupación nos establecementos hoteleiros. Grao de ocupación.
 - EOTRP_04: Ocupación nos establecementos de turismo rural. Pernoitas.

- EOTRVE_04: Ocupación nos establecementos de turismo rural. Viaxeiros entrados.
- EOTRGR_04: Ocupación nos establecementos de turismo rural. Grao de ocupación.
- Prezos e índices de prezos
 - CPM_15_XERAL_09: Índice de comercio polo miúdo (CNAE 2009). Índice xeral. Base 2015.
 - IPRI15XE: Índice de prezos industriais. Índice xeral. Base 2015.
- Demanda interna
 - MATTUR: Matriculación de vehículos. Turismos.
 - MATVCARGA: Matriculación de vehículos de carga.
- Sector exterior
 - CXM: Importacións con destino Galicia. Total.
 - CXX: Exportacións con orixe Galicia. Total.
 - IVU10_MTOT: Índices de valor unitario. Base 2010. Total importacións.
 - IVU10_XTOT: Índices de valor unitario. Base 2010. Total exportacións.
 - IVU10_MNE: Índices de valor unitario. Base 2010. Importacións non enerxéticas.

Os obxectivos do axuste estacional son identificar e eliminar as fluctuacións estacionais e os efectos de calendario que poidan enmascarar os movementos a curto e longo prazo nunha serie temporal e impedir unha comprensión clara dos fenómenos subxacentes. Expliquemos esta idea cun exemplo. Para isto, consideraremos a serie de datos de ocupación nos establecementos hoteleiros, contabilizando as pernoitas en Galicia, EOHP_04. Para facer o exemplo máis visual consideraremos só os datos comprendidos entre xaneiro de 1999 e xaneiro de 2001.

Comezaremos representando esta serie de datos inicial na Figura 1.1, onde podemos apreciar claramente certos comportamentos periódicos e repetitivos non desexables, en particular temos tres picos significativamente elevados na serie, que se producen nos meses de agosto de cada ano, mentres que nos primeiros meses do ano a serie toma valores moito menores. Este comportamento vén explicado polo feito de que o número de persoas que pernoctan en hoteis en Galicia é maior no verán (temporada alta). En particular, os meses de verán son os que teñen os picos más elevados, mentres que nos meses de inverno, en torno a xaneiro, é a época de menor ocupación hoteleira. Este fenómeno é o que definimos previamente como efecto estacional, onde podemos apreciar como temos un patrón que se repite ao longo dos anos. Ademais, temos outro efecto bastante visual que se produce á altura de marzo e abril, onde se aprecia un pequeno pico, este efecto é o asociado ao fenómeno de Semana Santa, xa que en Semana Santa a ocupación hoteleira tamén experimenta un aumento significativo do número de persoas que pernoctan en Galicia. Nótese que este efecto é moi importante a ter en conta xa que non podemos comparar un mes de marzo dun ano no que Semana Santa cae nese mes, con outro mes de marzo dun ano no que Semana Santa coincidiu en abril. Se seguimos falando dunha posible comparativa, podemos apreciar que no ano 2000 o pico de persoas que pernoctan en Galicia é inferior que o de 1999 ou 2001, pero antes de argumentar isto deberíamos ter en conta o número de días non laborables que ten cada ano nos distintos meses que o componen, xa que, por exemplo esa diferenza podería ser consecuencia de que o mes de agosto do ano 2000 ten un domingo menos que o ano 1999.

Todos estes efectos mencionados dificultan a análise da serie e a súa comparabilidade con outras series, xa que, a estacionalidade e o efecto de calendario están enmascarando o comportamento real da serie, por iso será de gran interese obter a serie correxida con axuste estacional, é dicir, a serie correxida do efecto estacional e de calendario, obtendo unha serie como a que vén representada na Figura 1.2 en cor azul. Agora podemos ver unha serie moito más suave con pequenas fluctuacións, onde os grandes picos e comportamentos periódicos da serie orixinal xa non están presentes.

O noso obxectivo tamén pode ser saber se a serie presenta tendencia á alza ou á baixa unha vez eliminadas as componentes propias da estación, xa que, eliminar ditas componentes debidas ao verán e á Semana Santa, entre

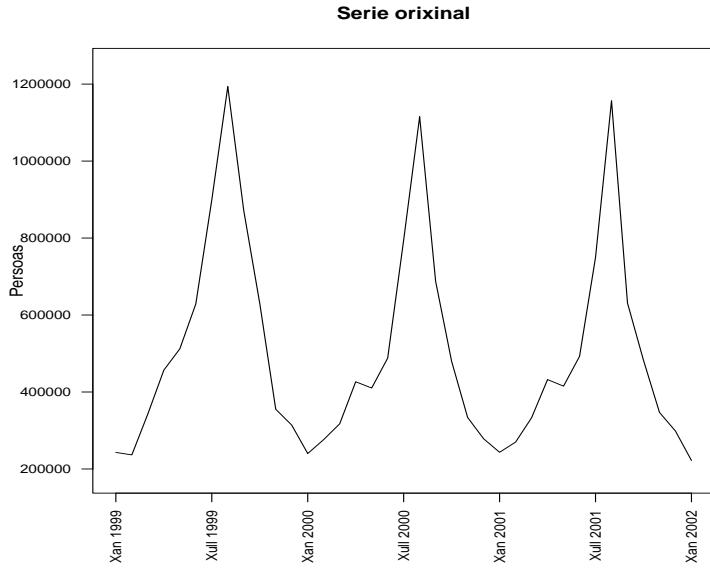


Figura 1.1: Serie orixinal de pernoitas nos establecementos hoteleiros.

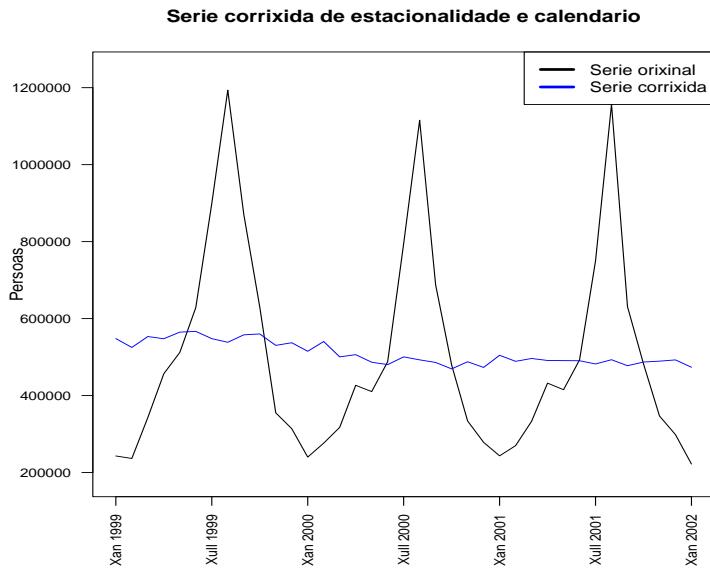


Figura 1.2: Serie orixinal de pernoitas nos establecementos hoteleiros xunto coa serie corrixida de estacionalidade e calendario, en cor azul.

outras, permítennos apreciar máis facilmente a presenza de tendencia (ou ciclotendencia). Consideremos de novo a mesma serie de ocupación nos establecementos hoteleiros de Galicia, contabilizando as pernoitas, pero agora cos datos comprendidos dende xaneiro de 1999 ata xaneiro de 2020 (xa que a tendencia é un fenómeno que se estuda a longo prazo). Na Figura 1.3 representamos en cor vermella a ciclotendencia da serie, acompañada da

serie orixinal e da serie corrixida de estacionalidade, que xa vimos previamente. Podemos apreciar tendencias á alza ou á baixa nos diferentes períodos comprendidos entre os xaneiros de 1999 e 2020, pero xa non temos os grandes picos da serie orixinal, que entorpecían a análise da serie levándonos a unha interpretación errónea do comportamento da tendencia da serie. Polo tanto, o axuste estacional é un proceso fundamental na interpretación de series temporais para a elaboración de políticas públicas.

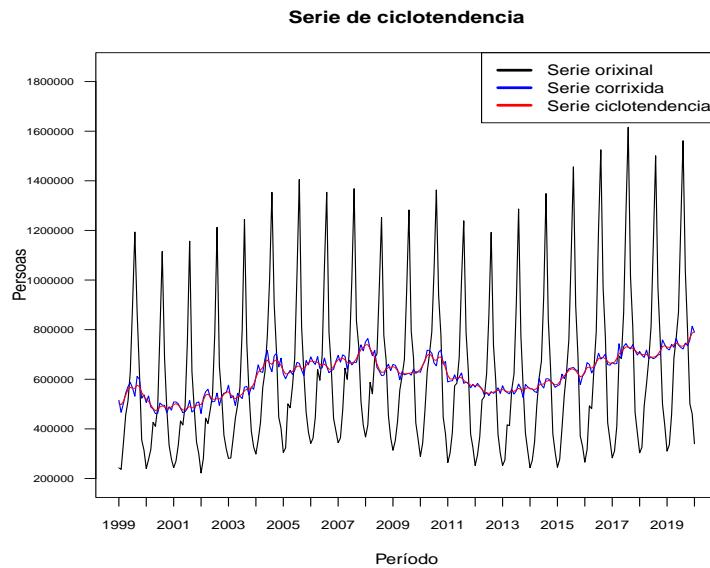


Figura 1.3: Serie orixinal de pernoitas nos establecementos hoteleiros, xunto coa serie correxida de estacionalidade e calendario, en cor azul, ademais da serie de tendencia en vermello.

O obxectivo deste traballo é facer unha revisión metodolóxica dos métodos empregados para facer os axustes estacionais e de calendario que agora mesmo están dispoñibles nas series seleccionadas do Banco de series de conxuntura. Proporanse e avaliaranse novos modelos de corrección das series e desenvolveranse algoritmos automáticos para o axuste e a grabación na base de datos.

Actualmente no IGE o procedemento para traballar coas series de conxuntura económica emprega o software TRAMO-SEATS. A continuación, daremos unha breve descripción da idea que se esconde detrás de cada un dos dous programas, TRAMO e SEATS, que deron lugar a dito software, desenvoltos por Victor Gómez e Agustín Maravall, desde o Banco de España (Gómez e Maravall, 1996). Por un lado, TRAMO (“Time Series Regression with ARIMA Noise, Missing Observations and Outliers”) é un programa escrito en Fortran que realiza estimacións, predicións e interpolacións de modelos de regresión con observacións faltantes e erros ARIMA, en presenza, polo xeral, de diferentes valores atípicos. Neste caso, o modelo ARIMA pódese identificar automáticamente. Polo outro lado, SEATS (“Signal Extraction in ARIMA Time Series”) é un programa escrito en Fortran que entra na clase dos chamados métodos baseados no modelo ARIMA para a descomposición dunha serie de tempo nas súas componentes non observadas (é dicir, para extraer dunha serie de tempo as súas diferentes sinais). O método concibiuse orixinariamente para o axuste estacional das series de tempo económicas, é dicir, para a eliminación do sinal estacional destas.

En contraste, no INE e Eurostat xa non traballan únicamente co software TRAMO-SEATS senón que pasaron a empregar o programa JDemetra+. A continuación describiremos brevemente os antecedentes que deron lugar á creación de JDemetra+.

Entre os métodos de axuste estacional que producen resultados más fiables para grandes conxuntos de da-

tos, os más empregados e recomendados son TRAMO-SEATS e X-13ARIMA-SEATS, considerados como os métodos de axuste estacional que producen mellores resultados. No caso deste último, X-13ARIMA-SEATS, desenvolto na Oficina do Censo de EE. UU. (U.S. Census Bureau, 2015), en linguaxe Fortran, tense que non é máis que o resultado de sucesivas melloras do software inicial X-11 da Oficina do Censo de EE. UU. (Shiskin et al., 1967). En particular, despois de X-11 apareceu o X-11-ARIMA do centro de estatística de Canadá (Dagum, 1980), aportando melloras sobre o X-11 en diversos aspectos, como foi a incorporación de novos elementos para a diagnose que proporcionaron mellores estimadores das compoñentes ao final da serie. A continuación formouse o programa X-12-ARIMA (U.S. Census Bureau, 2011) avantaxando ao X-11-ARIMA ao incorporar novos diagnósticos e algúns novos rasgos baseados na estimación de efectos especiais. Finalmente constitúese o software X-13ARIMA-SEATS como unha modificación do xa existente X-12-ARIMA que ofrece melloras no diagnóstico así como unha versión actualizada do software SEATS do Banco de España.

Ambos métodos, TRAMO-SEATS e X-13ARIMA-SEATS, divídense en dúas partes principais. A primeira parte, denominase preaxuste, e nela elimínanse os efectos deterministas da serie mediante un modelo de regresión cun erro modelado segundo un ARIMA. A segunda parte é a descomposición da serie temporal que ten como obxectivo estimar e eliminar unha compoñente estacional da serie temporal. TRAMO-SEATS e X-13ARIMA-SEATS utilizan un enfoque moi similar na primeira parte para estimar o mesmo modelo no paso de procesamento, pero difiren completamente no paso de descomposición. Polo tanto, comparar os resultados da descomposición a miúdo é difícil. Ademais, os seus diagnósticos enfócanse en diferentes aspectos e os seus resultados toman formas completamente diferentes.

Os dous programas de axuste estacional anteriores escribiríronse orixinalmente en Fortran, que actualmente se reconoce como unha linguaaxe de programación en declive. As limitacións de Fortran, especialmente para a creación de compoñentes reutilizables e para a xestión de problemas complexos, fan que o mantemento dos códigos sexa cada vez máis laborioso. Estos programas orixinarios de axuste estacional son comunmente percibidos polos usuarios como difíciles de operar. Polo tanto, para facilitar o acceso de persoas non especializadas no axuste estacional a estes métodos, Eurostat introduciu, en 2002, un novo software chamado Demetra. Ofrece unha interface fácil de usar para os dous algoritmos, TRAMO-SEATS e X-13ARIMA-SEATS, e facilita tamén a posible comparación dos resultados dos dous algoritmos. Aínda así, Demetra utiliza as bibliotecas Fortran, o que, xunto cun desenvolvemento de producto insuficiente e manexo de errores, é un factor que provocou unha rápida caída no uso do software.

En 2009, o Sistema Estatístico Europeo (ESS) lanzou as súas Directrices sobre o axuste estacional (“Guidelines on Seasonal Adjustment”, (Eurostat, 2009)) que tiñan como obxectivo harmonizar as prácticas europeas e mellorar a comparabilidade das estatísticas nacionais infraanuais, así como, conseguir unha mellora na calidade xeral da Unión Europea e dos agregados da zona euro. Como Demetra non podía adaptarse aos novos requisitos das Directrices, Eurostat, en cooperación co Banco Nacional de Bélxica (National Bank of Belgium, NBB), iniciou un proxecto, que foi lanzado posteriormente en 2012, cuxo obxectivo era desenvolver un software mellorado chamado Demetra+. Esta ferramenta proporciona un enfoque común para o axuste estacional utilizando os métodos TRAMO-SEATS e X-13ARIMA-SEATS, que é máis coherente coas Directrices. Inclúe unha interface gráfica unificada e diagnósticos de entrada e saída para os dous métodos. O código fonte de Demetra+ está escrito en C++ e utiliza os dous módulos Fortran orixinais, así como as bibliotecas .NET. En consecuencia, o software Demetra+ non é extensible e non se pode empregar en entornos de Tecnoloxía da Información (TI) que non sexan Windows. Por estas razóns, parece que a longo prazo non cumplirá coas expectativas dos usuarios.

Polo tanto, Eurostat tomou a iniciativa de crear un novo software baseado na experiencia de Demetra+ pero independiente da plataforma e extensible. O NBB estivo desenvolvendo desde 2012 o programa resultante, chamado JDemetra+ (Grudkowska, 2015). A nivel do usuario, en comparación con Demetra+, implementáronse numerosas melloras en JDemetra+, tanto en termos de deseño como de funcionalidades. Pero a innovación máis crítica é a reescritura dos códigos Fortran orixinais de X-13ARIMA-SEATS e TRAMO-SEATS en linguaxe Java, segundo un enfoque real orientado a obxectos. En particular, na Figura 1.4 podemos ver de forma esquemática a descomposición nos distintos programas que foron xurdindo ao longo dos anos, como acabamos de comentar, e acabaron dando lugar á creación do programa JDemetra+.

Neste traballo centrarémonos na utilización do paquete RJDemetra+ que non é máis que unha adaptación

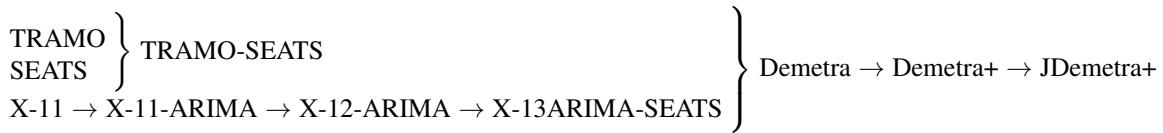


Figura 1.4: Esquema da estrutura subxacente baixo o programa JDemetra+.

do software de axuste estacional JDemetra+ para R Project (entorno e linguaxe de programación orientada á análise estadística). En particular, ofrece acceso á maioría das opcións e saídas de JDemetra+, incluídos os dous principais métodos de axuste estacional TRAMO-SEATS e X-13ARIMA-SEATS.

Dende o INE establecense unha serie de recomendacións para o tratamento de series de convivencia económica seguindo as indicacións de Eurostat (consultar Instituto Nacional de Estadística, 2019).

Comezaremos falando do axuste de calendario, que recomenda considerar un modelo de regresión cun erro modelado segundo un ARIMA, aplicando contrastes de significación dos efectos. Entre os efectos de calendario considéranse un efecto asociado á Semana Santa (que pode estar formado por un ou máis regresores), un efecto do ano bisesto e un efecto que contabilice o número de días hábiles de cada mes (recoméndase contrastar se existen diferenzas entre os distintos días hábiles).

Por outro lado, o axuste estacional do INE baséase en modelos ARIMA. En particular, nos casos nos que a serie é inferior a 7 anos non se recomenda o axuste estacional.

En canto á selección do modelo, considérase un proceso automático de selección e realizanse contrastes estadísticos e análise dos seus residuos para avaliar se o modelo é adecuado. Ademais no INE incídese no feito de que para series problemáticas o modelo se debe construir a man. A posterior validación do modelo axustado estacionalmente desenvolverase mediante análise gráfica e criterios descriptivos paramétricos e non paramétricos.

Con respecto ao enfoque de tratamiento das series, o INE diferencia entre enfoque directo (todas as series axústanse estacionalmente de forma individual) ou indirecto (algunha serie obtense como combinación lineal de dúas ou máis series axustadas de efecto estacional). Na práctica téndese a empregar o método directo por ser o más simple.

Tamén debemos ter en conta que os valores da compoñente de tendencia ao final da serie son estimados por extrapolación, e polo tanto os valores estimados para esta compoñente nos datos máis recentes son moi incertos e poden sufrir problemas de cambio de fase.

Por último, é necesario facer mención ao problema engadido que supoñen os valores atípicos. Definiremos un valor atípico, ou outlier, como unha observación anómala en comparación co resto de datos contidos nun determinado conxunto de datos. Pode deberse a diversos motivos, dende erros humanos ata causas como que o dato proceda dunha poboación diferente da que se pretende estudiar. A presenza de valores atípicos á hora de modelar unha serie de tempo pode levarnos a obter modelos resultantes do axuste insatisfactorios, debido á influencia que exercen ditos valores e en consecuencia, debén ser tidos en conta para que non entorpezan o proceso de estimación do modelo. Ademais, faise especial fincapé na dificultade implícita que supón modelar este tipo de datos se se atopan ao final da serie, onde unha mala interpretación, pode levar a perturbar considerablemente a serie axustada. Estes valores identifícanse de forma que os debidos a erros deben ser corregidos das series brutas, e o resto de atípicos que teñan unha clara explicación (información coñecida) engadidos como regresores. Como alternativa poderíase considerar o proceso totalmente automático para a detección de atípicos. Debemos facer mención do caso particular da aparición do COVID-19. Dende Europa, (véxase Eurostat, 2020), recoméndase tratar estos datos como un punto ao final da serie que se modela co tratamento aplicado para atípicos ao final dunha serie ou ao comezo dun gran cambio económico. Comezarán modelándose como un atípico aditivo (ou cambio puntual) ao final da serie e non se poderán cambiar por outro tipo de atípico (cambio transitorio ou de nivel) mentres non disponhamos de máis información.

A continuación, estructuraremos o traballo en catro partes. A primeira parte (Capítulo 2) estará dedicada a introducir brevemente as ideas principais da metodoloxía implementada internamente en JDemetra+. Deseguido, no Capítulo 3, realizaremos o axuste estacional das series de conxuntura de Galicia co programa JDemetra+. Co obxectivo de obter un proceso automático, no Capítulo 4 realizarase unha adaptación do procedemento visto para o axuste estacional en JDemetra+ coa linguaxe de programación R, obtendo un código coa implementación automática do proceso. Ademais, neste apartado, adaptaremos as series de tempo correxidas resultantes para a súa posterior introdución na Base de datos do IGE, de modo que na web do IGE se substuirán os antigos modelos de axuste estacional polos obtidos co procedemento automático implementado en R. Por último, no Capítulo 5, recompílanse as conclusións de todo o procedemento levado a cabo así como as posibles melloras futuras. Ademais temos dispoñibles dous apéndices, dos cales, o Apéndice A recolle os parámetros finais dos modelos de axuste estacional óptimos obtidos para cada unha das series seleccionadas de entre as consideradas más relevantes do Banco de series de conxuntura. Pola súa parte, o Apéndice B recolle o código de R de automatización do proceso de axuste estacional que queda dispoñible e en uso no IGE, para a corrección das series de tempo.

Capítulo 2

Metodoloxía

Todo o procedemento levado a cabo para modelizar as series de tempo e eliminarlle o efecto estacional explicarase a continuación seguindo as pautas que veñen recollidas no apéndice do libro de referencia do axuste estacional de Eurostat, que se pode consultar en Grudkowska (2016).

Antes de dar unha definición formal do concepto de serie de tempo debemos introducir unha serie de conceptos que facilitarán a súa comprensión, para isto apoiamos en Peña (2005).

Comezaremos por definir un proceso estocástico, que dende un punto de vista intuitivo, non é máis que unha secuencia de datos que evolucionan no tempo. Formalmente, denotaremos un proceso estocástico como un conxunto de variables aleatorias $\{Y_t\}_{t \in C}$, definidas sobre o mesmo espacio de probabilidade, onde C denota un conxunto de números como pode ser o conxunto dos números enteiros (\mathbb{Z}), o dos números naturais (\mathbb{N}), etc. En particular, podemos considerar o proceso estocástico $\{Y_t\}_{t \in \mathbb{Z}}$, de tal modo que o noso proceso estocástico seguirá a estrutura

$$\dots, Y_{-2}, Y_{-1}, Y_0, Y_1, Y_2, \dots$$

onde o subíndice t de cada variable aleatoria representa o instante de tempo no que se observa.

Unha observación do proceso estocástico será da forma

$$\dots, y_{-2}, y_{-1}, y_0, y_1, y_2, \dots$$

e coñécese como unha realización ou traxectoria do mesmo. Logo, se consideramos unha serie de tempo dada por y_1, y_2, \dots, y_T , con T denotando o número total de elementos da serie, teremos unha realización ou traxectoria parcial dun proceso estocástico. Noutras palabras, as series temporais representan un caso particular dos procesos estocásticos.

Dado un proceso estocástico $\{Y_t\}_{t \in \mathbb{N}}$, defínense:

- Función de medias: $\mu_t = \mathbb{E}(Y_t)$, medida de posición de carácter central de Y_t .
- Función de varianzas: $\sigma_t^2 = \text{Var}(Y_t) = \mathbb{E}((Y_t - \mu_t)^2)$, medida do grao de variabilidade de Y_t .
- Función de autocovarianzas: $\gamma(s, t) = \text{Cov}(Y_s, Y_t) = \mathbb{E}((Y_s - \mu_s)(Y_t - \mu_t))$, medida do grao de dependencia lineal existente entre Y_s e Y_t .
- Función de autocorrelacións simples (fas):

$$\rho(s, t) = \frac{\gamma(s, t)}{\sigma_s \sigma_t},$$

medida do grao de dependencia lineal existente entre Y_s e Y_t . Toma valores en $[-1, 1]$.

- Función de autocorrelacións parciais (fap):

$$\alpha(s, t) = \frac{Cov(Y_s - \hat{Y}_s^{(s,t)}, Y_t - \hat{Y}_t^{(s,t)})}{\sqrt{Var(Y_s - \hat{Y}_s^{(s,t)})} Var(Y_t - \hat{Y}_t^{(s,t)})},$$

onde $\hat{Y}_j^{(s,t)}$ denota ó mellor predictor lineal de Y_j construído a partir das variables medidas nos instantes comprendidos entre s e t (sen considerar s nin t). $\alpha(s, t)$ é unha medida do grao de dependencia lineal existente entre Y_s e Y_t , unha vez que se lles subtraeu o efecto lineal que exercen sobre cada unha delas as variables medidas nos instantes comprendidos entre s e t . Toma valores en $[-1, 1]$.

Un proceso estocástico $\{Y_t\}_{t \in \mathbb{N}}$ dise que é estacionario se se verifica:

1. $\mu_t = \mu \forall t$.
2. $\sigma_t^2 = \sigma^2, \forall t$.
3. $\gamma(t, t+k) = \gamma_k, \forall t, k$.

É dicir, a estacionariedade dota ao proceso estocástico de propiedades de estabilidade en media, varianza e autocovarianzas.

Polo tanto, estamos en condicións de definir unha serie de tempo, y_t , como unha realización dun proceso estocástico. Calquera proceso estocástico estacionario en covarianza pode representarse mediante a chamada Descomposición de Wold,

$$y_t = \mu_t + x_t, \quad (2.1)$$

onde μ_t é unha compoñente determinista (como pode ser unha constante, unha función do tempo, etc) e x_t é unha compoñente non determinista, que se pode expresar da seguinte forma:

$$x_t = \sum_{j=0}^{\infty} \psi_j a_{t-j},$$

onde $\psi_0 = 1$, $\sum_{j=0}^{\infty} \psi_j^2 < \infty$ e a_t é o erro dado por ruído branco. Enténdese por ruído branco a unha colección de variables aleatorias incorreladas con media cero e varianza constante, V_a . En particular, o termo de erro, a_t , representa o erro de predición dun período por diante de x_t , é dicir,

$$a_t = x_t - \hat{x}_{t-1}, \quad (2.2)$$

onde \hat{x}_{t-1} é a predición de x_t realizada no período $t-1$. Como a_t representa o que hai de novo en x_t no punto t , é dicir, que non estaba contido nos valores pasados de x_t , tamén se denomina innovación do proceso.

A Descomposición de Wold presenta un proceso estocástico como a suma dunha compoñente linealmente determinista μ_t e dunha compoñente linealmente non determinística $\sum_{j=0}^{\infty} \psi_j a_{t-j}$, a primeira é perfectamente predecible unha vez que se coñece a historia do proceso y_{t-1} e o segundo é imposible de predecir perfectamente. Isto explica que o proceso estocástico non se poida predecir sen cometer un certo erro.

Por outro lado sexan $\{a_t\}_t$ un proceso de ruído branco e $\{Y_t\}_{t \in \mathbb{N}}$ un proceso estocástico. Defínese un proceso lineal como aquel que admite unha representación do tipo

$$Y_t = c + \sum_{j=-\infty}^{\infty} \psi_j a_{t-j}$$

sendo c unha constante, con $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$. Os procesos lineais son estacionarios. En particular, se na Descomposición de Wold eliminamos a compoñente determinista, quedando soamente coa compoñente non

determinista, x_t , temos un proceso lineal, e polo tanto estacionario. En consecuencia, calquera proceso estacionario pode ser presentado como un proceso lineal.

A continuación introducirémonos no procedemento levado a cabo para o axuste das series de tempo, comezando por presentar a estrutura dos modelos ARIMA (véxase Brockwell e Davis, 2002), que constitúen unha parte fundamental do proceso de axuste das series de tempo.

Unha serie de tempo x_t , con $t \in \mathbb{N}$, pódese modelar como un proceso ARIMA seguindo a estrutura

$$\phi(B)\delta(B)x_t = \theta(B)a_t,$$

sendo $\phi(B)$, $\theta(B)$ e $\delta(B)$ polinomios finitos en B e a_t a compoñente de ruído branco.

$\phi(B)$ é un polinomio estacionario autorregresivo (AR) en B , resultado do producto do polinomio AR regular estacionario en B ($\phi_p(B)$) e o polinomio AR estacional estacionario en B^s ($\Phi_p(B^s)$):

$$\phi(B) = \phi_p(B)\Phi_p(B^s) = (1 + \phi_1B + \dots + \phi_pB^p)(1 + \Phi_1B^s + \dots + \Phi_pB^{Ps})$$

onde

- p é o número de termos AR regulares (en JDemetra+ $p \leq 3$).
- P é o número de termos AR estacionais (en JDemetra+ $P \leq 1$).
- s é o número de observacións ao ano (frecuencia da serie temporal).

$\theta(B)$ é un polinomio de medias móbiles invertible (MA) en B , que é producto do polinomio MA regular invertible en B ($\theta_q(B)$) e o polinomio MA estacional invertible en B ($\Theta_q(B^s)$):

$$\theta(B) = \theta_q(b)\Theta_Q(B^s) = (1 + \theta_1B + \dots + \theta_qB^q)(1 + \Theta_1B^s + \dots + \Theta_QB^{Qs}),$$

onde:

- q é o número de termos MA regulares (en JDemetra+ $q \leq 3$).
- Q é o número de termos MA estacionais (en JDemetra+ $Q \leq 1$).

$\delta(B)$ é o polinomio de diferenciación regular e estacional, que se emprega para correxir a non estacionariedade en media da serie.

$$\delta(B) = (1 - B)^d(1 - B^s)^D,$$

onde:

- d é a orde de diferenciación regular (en JDemetra+ $d \leq 2$).
- D é a orde de diferenciación estacional (en JDemetra+ $D \leq 1$).

2.1. Linealización

En xeral, para as series de tempo observadas, os supostos relativos á estacionariedade do proceso, é dicir estabilidade en media, varianza e covarianza, non se manteñen por varios motivos. En primeiro lugar, a variación da serie temporal pode cambiar no tempo. En segundo lugar, a maioría das series cronolóxicas observadas presentan unha media que non se pode asumir constante debido á presenza dunha tendencia e dos movementos estacionais. Finalmente, as series temporais observadas adoitan conter efectos deterministas como son os valores atípicos e os efectos de calendario. Polo tanto, na práctica cómpre aplicar un axuste previo á serie temporal para solventar ditos inconvenientes e conseguir a estacionariedade da serie.

2.1.1. Estabilidade en media e varianza

Comezaremos buscando series estacionarias tanto en media como en varianza. A estacionariedade en media conséguese aplicando diferencias regulares (d) e de efecto estacional (D), segundo corresponda, que se poden modelar como segue:

$$\delta(B)y_t = (1 - B)^d(1 - B^s)^D y_t,$$

onde B denota o operador retardo regular definido por $By_t = y_{t-1}$ e B^s o operador retardo estacional, dado por $B^s y_t = y_{t-s}$. Á súa vez, d e D denotan as ordes de diferenciación regular e estacional que correspondan en cada caso.

Para conseguir estacionariedade en varianza, é dicir, eliminar a heterocedasticidade, aplícase a seguinte transformación denotada como transformación Box-Cox, para máis información consultar Box e Cox (1964), que modifica a serie y_t como segue,

$$\begin{cases} \frac{(y_t+c)^\lambda - 1}{\lambda} & \text{se } \lambda \neq 0, \\ \log(y_t + c) & \text{se } \lambda = 0. \end{cases} \quad (2.3)$$

sendo c unha constante que verifique $(y_t + c) > 0 \forall t$. Este procedemento consegue estabilizar a varianza dos valores transformados da serie temporal y_t .

2.1.2. Eliminación dos efectos deterministas

Por último, temos que encargarnos da corrección dos efectos deterministas, que veñen dados polos valores atípicos e polos efectos de calendario.

Datos atípicos

Un dato atípico ou outlier, como xa adiantábamos no Capítulo 1, é unha observación anómala en comparación co resto de datos contidos nun determinado conxunto de observacións. Podemos expresar o seu impacto sobre as series observadas como:

$$y_t = \sum_{j=1}^k \xi_j(B) w_j I_t^{(\tau_j)} + x_t,$$

onde

- y_t denota a serie de tempo observada.
- x_t é unha serie que segue o modelo ARIMA.
- w_j é un impacto inicial do valor atípico en tempo $t = \tau_j$.
- $I_t^{(\tau_j)}$ é unha variable indicadora tal que vale 1 para $t = \tau_j$ e 0 en caso contrario.
- $\xi_j(B)$ é unha expresión que determina a dinámica do valor atípico que ocorre en tempo $t = \tau_j$.
- B é o operador retardo (é dicir, $B^k x_t = x_{t-k}$).

Podemos diferenciar catro tipos de atípicos:

- Atípico aditivo (AO): un punto atípico que ocorre nun momento dado t_0 . Para o valor atípico aditivo tense $\xi_j(B) = 1$, que resulta na variable de regresión:

$$\text{AO}_t^{t_0} = \begin{cases} 1 & \text{para } t = t_0, \\ 0 & \text{para } t \neq t_0. \end{cases}$$

- Cambio de nivel (LS): variable para un cambio de nivel constante que comeza no momento dado t_0 . Para o cambio de nivel tense a función de dinámica $\xi_j(B) = \frac{1}{1-B}$, que resulta na variable de regresión:

$$\text{LS}_t^{t_0} = \begin{cases} 0 & \text{para } t < t_0, \\ 1 & \text{para } t \geq t_0. \end{cases}$$

- Cambio temporal (TC): variable para un efecto sobre o tempo dado t_0 que decae exponencialmente durante os seguintes períodos. Para o cambio temporal tense a función de dinámica $\xi_j(B) = \frac{1}{1-\delta B}$, que resulta na variable de regresión:

$$\text{TC}_t^{t_0} = \begin{cases} 0 & \text{para } t < t_0, \\ \delta^{t-t_0} & \text{para } t \geq t_0. \end{cases}$$

onde δ é unha taxa de decrecemento cara ó nivel anterior ($0 < \delta < 1$). Nótese que cando δ tende a 0, o TC redúcese a un AO, mentres que se δ tende a 1, o TC redúcese a un LS.

- Atípico estacional (SO): variable que recolle un cambio brusco no patrón estacional no tempo dado t_0 e mantén o nivel da serie cun cambio significativo repartido nos períodos restantes. Está modelado pola variable de regresión:

$$\text{SO}_t^{t_0} = \begin{cases} 0 & \text{para } t < t_0, \\ 1 & \text{para } t \geq t_0, t \text{ no mesmo mes/trimestre que } t_0, \\ \frac{-1}{s-1} & \text{noutro caso,} \end{cases}$$

onde s é a frecuencia da serie temporal, é dicir, $s = 12$ para unha serie mensual ou $s = 4$ para unha serie trimestral.

De forma máis intuitiva, un AO representa un pico illado na serie, un LS unha función escalonada, un TC un pico que tarda uns períodos en desaparecer e un SO varios picos que se producen de forma repetitiva.

Os valores atípicos, usados con frecuencia na modelización de series temporais económicas estacionais, poden ser detectados automaticamente por JDemetra+. Con todo, cabe mencionar que algunas posibles combinacións destes atípicos non son coherentes polo que deben evitarse. Como por exemplo, considerar varios AO seguidos dun TC ou dun LS, non tería sentido, aínda que se obtivese un axuste razoable.

Efectos de calendario

Por outra parte, tamén debemos ter en conta os efectos de calendario. Podemos definir estes efectos como a parte do movemento dunha serie temporal causada polo diferente número de cada tipo de días da semana nos distintos meses ou trimestres. A importancia de considerar os efectos de calendario recae no feito de que o número de ocorrencias dun determinado día da semana nun mes, ou trimestre, difire dun ano a outro.

Para facilitar a súa comprensión podemos poñer un par de exemplos sinxelos do efecto de calendario. Para o caso particular de Galicia as festividades de San Xosé (19 de marzo) e San Xoán (24 de xuño) non son consideradas como días festivos todos os anos, de modo que non son comparables dous anos nos que esta festividade ten consideracións diferentes. Ademais, temos o efecto de que un festivo coincide en día de semana ou fin de semana, pois no segundo dos casos o efecto do festivo xa non vai ser tan significativo. Tamén se debe mencionar o efecto de calendario no sentido visto no Capítulo 1, onde había claras diferencias na serie entre as estacións do ano, xa que en certos sectores como o do turismo a súa actividade increméntase substancialmente no verán. Por outro lado, o efecto de Semana Santa sempre produce alteracións na serie que deben ser tidas en

conta, alteracións que a súa vez son móbiles xa que a Semana Santa non coincide sempre nas mesmas datas, nin sequera no mesmo mes. Tamén comentamos de antemán que a variación do número de días da semana e fin de semana dun mes con respecto a outro, aínda que sexa pequena, ten un efecto sobre a serie. Polo tanto, todos estes acontecementos asociados ao calendario deben ser correxidos para que non enmascaren o verdadeiro comportamento da serie de tempo.

Un xeito natural de modelar os efectos de calendario consiste en distribuír os días correspondentes a cada período en diferentes grupos, cada un dos cales será representado no modelo engadindo unha variable de regresión asociada. Deste modo, cada variable de regresión virá definida polo número de días que contén para cada período. As dúas clasificacións máis empregadas son:

- Trading days: xéranse 7 regresores ou grupos, 6 asociados aos efectos de cada día da semana restándolle o efecto do domingo e, un último coa lonxitude do período, que en particular recolle o efecto do ano bisesto. O regresor dos luns calcúllase como o número de luns nun mes (ou trimestre) menos o número de domingos dese mes (ou trimestre). De xeito análogo constrúense os regresores de todos os días da semana salvo o domingo. Tamén se xera un regresor que contén o número de días de cada mes (trimestre) que capture o efecto do ano bisesto.

De forma explícita, o regresor de Trading days e ano bisesto quedaría como:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} M \\ T \\ W \\ Th \\ F \\ Sat \\ S \end{bmatrix} = \begin{bmatrix} M - S \\ T - S \\ W - S \\ Th - S \\ F - S \\ Sat - S \\ \text{Length of period} \end{bmatrix}$$

- Working days: xéranse 2 grupos, un para contrastar o efecto dos días da semana cos do fin de semana e outro coa lonxitude do período, que de novo recolle o efecto do ano bisesto. O regresor para a diferenza entre os días da semana (de luns a venres) e os do fin de semana (sábados e domingos) costrúese como o número de días da semana nun mes (ou trimestre) menos o número de días de fin de semana dese mes (ou trimestre) multiplicados por $5/2$, onde a fracción $5/2$ serve para homoxeneizar os dous elementos da diferenza.

De forma explícita, o regresor de Working days e ano bisesto quedaría como:

$$\begin{bmatrix} 1 & -5/2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \text{Week} \\ \text{Weekend} \end{bmatrix} = \begin{bmatrix} \text{Contrast week} \\ \text{Length of period} \end{bmatrix}$$

Por defecto, o calendario que se considera en JDmetra+ correspondece coas variables habituais de xornada laboral en función do Calendario Gregoriano, coa posibilidade de engadir algúns festivos específicos. Ditos festivos trátanse como domingos e inclúense no grupo correspondente aos días non hábiles. Este enfoque supón que a actividade económica dos festivos nacionais é o mesmo (ou moi próximo) ao nivel de actividade típico dos

domingos. Polo tanto, agás se o festivo coincide en domingo, a aparición dun festivo leva a corrección en dous grupos, é dicir, no grupo que contén o día da semana no que cae festivo e no grupo que contén os domingos.

JDemetra+ non contén os festivos de cada país, por defecto, só considera os sábados e domingos como días inhábiles, polo que, para considerar os festivos concretos de cada rexión debemos crear o noso propio calendario. Hai tres opcións de calendario dispoñibles:

- National calendars: é apropiado para definir un calendario que inclúa festivos nacionais específico.
- Chained calendars: definidos por dous calendarios nacionais e unha data de corte entre eles.
- Composite calendars: definidos como unha suma ponderada de diferentes calendarios nacionais.

En particular para modelar o noso propio calendario, temos implementadas catro tipos de festividades diferentes:

- Días fixos (Fixed): define un festivo que se celebra sempre nun día concreto do ano, que sempre se produce o mesmo día do mes, por exemplo, o 1 de xaneiro.
- Días relacionados coa Pascua (Easter related): indica un festivo cuxa data depende da Semana Santa. Prodúcese con referencia ao día de Pascua $\pm n$ días, por exemplo, Pentecostés.
- Días fixos da semana (Fixed week): esta opción crea un festivo fixo que siempre cae nunha semana específica dun determinado mes. Por exemplo, o Día do Traballo celebrado en USA o primeiro luns de setembro.
- Día distinguido (Special Day): permite ao usuario escoller un día festivo dunha lista predefinida de festivos, que inclúe os festivos móbiles e fixos máis populares. Por exemplo, algunas das datas que veñen recollidas son o Nadal, a Ascensión e Halloween, entre outros.

O efecto de calendario pódese dividir en tres compoñentes. Para comprender mellor esta idea faremos un pequeno inciso explicando a descomposición aplicable de forma xeral para unha serie de tempo.

As series temporais estacionarias, X_t , pódense descompoñer en tres compoñentes principais, cada unha das cales representa o impacto de certos tipos de fenómenos na serie, que son as seguintes:

- A ciclo-tendencia (T_t) que capture o comportamento a longo e medio prazo. En xeral, a ciclo-tendencia consta de dous subcompoñentes:
 1. A tendencia ou evolución a longo prazo da serie. Comprende un período superior aos 8 anos.
 2. O ciclo, que representa o movemento suave, casi periódico, en torno á evolución a longo prazo da serie. Comprende un período de entre 2 e 8 anos.

A efectos de axuste estacional, non se separan a tendencia a longo prazo do ciclo, xa que estas dous compoñentes normalmente son demasiado curtas para realizar unha estimación fiable. En consecuencia, como xa adiantábamos no Capítulo 1, estímase únicamente unha compoñente para o ciclo e a tendencia, á que nos referiremos como tendencia (T_t).

- A compoñente estacional (S_t) que representa as fluctuacións intraanuais, mensuais ou trimestrais, que se repiten con maior ou menor regularidade ano tras ano.
- A compoñente irregular (I_t) combinando todas as demais fluctuacións más ou menos erráticas non cubertas polas compoñentes anteriores.

Pódense considerar dous modelos de descomposición, un aditivo, $X_t = T_t + S_t + I_t$, e outro multiplicativo, $X_t = T_t \cdot S_t \cdot I_t$, onde o modelo multiplicativo obtense de forma equivalente ao aditivo, simplemente aplicando logaritmos sobre a serie.

Entón, volvendo ao efecto de calendario, tense que dito efecto se pode dividir nun efecto medio, unha parte estacional e unha parte estrutural. O efecto da media é independente do período e debe estar asociado coa componente de tendencia. A parte estacional xorde das propiedades do calendario que se repiten cada ano, por exemplo, o número de días laborables en meses de 31 días son, de media, máis que en meses de 30 días. Este efecto é captado pola componente estacional (excepto o efecto do ano bisesto). Por último, a parte estrutural do efecto calendario é a que pertence aos efectos deterministas, e polo tanto, a que nos interesa modelar no axuste de calendario. Dito efecto parcial pódese estimar centrando o regresor do calendario, é dicir, eliminando a súa media mensual/trimestral a longo prazo.

Conxuntamente, o proceso de eliminación das componentes deterministas da serie de tempo orixinal pódese representar como un modelo de regresión con erro ARIMA, da forma

$$y_t = z_t \beta + x_t, \quad (2.4)$$

onde

- y_t é a serie orixinal.
- $\beta = (\beta_1, \dots, \beta_n)$ un vector de coeficientes de regresión.
- $z_t = (z_{1t}, \dots, z_{nt})$ n variables de regresión (as variables do efecto dos días laborais, o efecto do ano bisesto, o efecto de semana santa, valores atípicos,...).
- x_t unha perturbación ou componente de erro que segue o proceso xeral ARIMA, que por defecto, identifícarase de modo automático.

Todo o procedemento levado a cabo para acadar a estacionariedade da serie denomináñase conviudamente como preaxuste ou linealización e pódense realizar cos modelos TRAMO ou RegARIMA, en función de se estamos a empregar a implementación de TRAMO-SEATS ou a de X-13ARIMA-SEATS, respectivamente. A estimación realizada con ditos modelos proporciona a parte estocástica da serie temporal, chamada serie linealizada, que se supón que é unha realización dun proceso estocástico lineal.

O obxectivo principal do axuste estacional é eliminar a componente estacional non observable da serie. Ambos os métodos de axuste estacional aquí analizados inclúen un paso de modelado que ten como obxectivo estudar o desenvolvemento da serie temporal e proporcionar unha mellor entrada para a posterior descomposición en componentes, onde a ferramenta que se utiliza con frecuencia para este fin é o modelo ARIMA. Non obstante, a presenza de efectos deterministas impiden a identificación dun modelo ARIMA óptimo debido ao importante sesgo na estimación dos parámetros das funcións de autocorrelación da mostra, tanto global como parcial. Polo tanto, antes de nada, a serie orixinal debe ser corrixida por calquera efecto determinista, é dicir, debemos comezar realizando un proceso de linealización como o que acabamos de describir, obtendo a posteriori unha serie que si poderá ser modelada por ARIMA.

2.1.3. Identificación automática do modelo ARIMA

Os modelos ARIMA, que serán fundamentais na modelización da serie de tempo, pódense obter de forma automática. En función de se estamos a traballar co programa TRAMO-SEATS ou X-13ARIMA-SEATS o procedemento de identificación automática de modelos ARIMA presenta certas variacións.

TRAMO

Por un lado, o algoritmo para a identificación automática de modelos implementado en TRAMO pódese estructurar en dúas etapas iterativas: a primeira consiste na detección e corrección automática de valores atípicos e efectos de calendario e a segunda na identificación automática do modelo.

Paso 1 Modelo predeterminado de liñas aéreas e test de transformación Box-Cox.

En primeiro lugar, axústanse á serie temporal dous modelos de liñas aéreas (para máis información sobre estes modelos véxase Box e Jenkins, 1976), que son un tipo de modelos ARIMA que presentan a seguinte estrutura:

$$\text{ARIMA } (0, 1, 1)(0, 1, 1)$$

Considérase un en logarítmos ($\lambda = 0$) e outro sen logarítmos ($\lambda = 1$), de xeito que se poida realizar un test de comparación entre eles. É dicir, o primeiro paso é a realización dun test para decidir se se aplica a transformación Box-Cox para correxir a heterocedasticidade da serie. Por defecto, tense un lixeiro sesgo ou tendencia cara á transformación logarítmica.

Paso 2 Efectos de calendario e atípicos.

Neste paso realiza un test sobre os efectos de calendario utilizando o modelo ARIMA (obtido no paso previo) para xerar a compoñente de erro da ecuación (2.4) e, se posteriormente se modifica o modelo, repetiríase dito test empregando o novo modelo. Nótese que para as series estacionais o modelo predeterminado é o modelo de liñas aéreas, mentres que para as series non estacionais utilízase o modelo ARIMA(0, 1, 1)(0, 0, 0) cun termo constante. Polo tanto, o test sobre a posible presenza de estacionalidade determinará o modelo predeterminado que se usará nos seguintes pasos.

Unha vez completadas estas probas previas, corríxese a serie orixinal para todos os valores atípicos preespecificados e os efectos de calendario.

Paso 3 Axuste dun modelo ARIMA para a serie correxida de efectos deterministas ($x_t = y_t - z_t\beta$).

A continuación, identifícase a orde do polinomio diferenciador $\delta(B)$, encargado de estabilizar a media, e decídese se se especifica ou non unha media para a serie. A identificación da orde dos polinomios restantes do modelo ARIMA, é dicir, $\phi(B)$ e $\theta(B)$, realiza por minimización do Criterio de Información Bayesiano (BIC), con algunas restriccións destinadas a aumentar a parsimonia e favorecer modelos equilibrados. Recordar que os modelos parsimoniosos son aqueles que teñen un gran poder explicativo utilizando un número relativamente pequeno de parámetros, mentres que os modelos equilibrados son aqueles nos que a orde dos operadores AR e diferenciador, combinados, son iguais á orde do operador MA. Este procedemento xera os valores iniciais dos parámetros do modelo ARIMA.

Paso 4 Eliminación dos efectos deterministas do novo modelo ARIMA obtido.

Os efectos do calendario que foron identificados no modelo predeterminado, inclúense no novo modelo ARIMA sempre que estes efectos continuen sendo significativos para o novo modelo.

No procedemento de detección de valores atípicos utilízanse os residuos estimados do modelo ARIMA obtivo no paso anterior, onde, para cada observación compútanse tests para todos os tipos de valores atípicos considerados no procedemento automático (AO, TC, LS, SO). Se hai varios valores atípicos para os que o valor absoluto do estadístico é maior que un valor crítico, selecciónase aquel que teña o maior valor absoluto. Despois de corrixir o valor atípico identificado, iníciase de novo o proceso para comprobar se hai outro valor atípico e este procedemento repítense ata que, para ningún dos valores atípicos potenciais, o estadístico supere o valor crítico.

Paso 5 Comparación de modelos obtidos ao longo do proceso iterativo seleccionando o mellor.

Por último, compróbase a adecuación dos residuos estimados do modelo ARIMA final, correxido de atípicos e efectos de calendario, fronte aos residuos estimados producidos polo modelo equilibrado (que foi previamente seleccionado polo criterio BIC). Como é de esperar, o modelo final identificado debe mostrar algunha mellora con respecto ao modelo predeterminado (de liñas aéreas) nestes diagnósticos residuais, xa que, en caso contrario, o programa aceptará o modelo predeterminado.

RegARIMA

Por outro lado, se agora consideramos o algoritmo RegARIMA tense que en esencia é moi semellante ao procedemento presentado en TRAMO, con lixeiras modificacións propias do software X-13ARIMA-SEATS. En particular, os valores dos parámetros predeterminados son diferentes aos empregados en TRAMO. Ademais, para escoller os parámetros do modelo ARIMA, o RegARIMA utiliza o Criterio de Información de Akaike corrixido (AICc) mentres que TRAMO utilizaba BIC con algunas restricións destinadas a aumentar a parsimonia e favorecer modelos equilibrados. De feito o test de transformación logarítmica en RegARIMA tamén está baseado no AICc, en vez de na estimación por máxima verosimilitude do parámetro da transformación Box-Cox, que era o procedemento de TRAMO. Máis exemplos de extensións que son propias unicamente de RegARIMA son o tratamento especial do efecto dos anos bisestos no modelo multiplicativo, así como, a detección automática da duración do efecto de Semana Santa, ou a opción de reducir unha serie de cambios de nivel (LS) a un cambio de nivel temporal (TC). Ademais, por defecto, RegARIMA non favorece os modelos equilibrados. En consonancia con todo isto, o modelo seleccionado por RegARIMA pode diferir do modelo que seleccionaría TRAMO.

Criterios de selección de modelos

Vimos de mencionar os criterios BIC e AICc como criterios de selección de modelos, servindo de ferramentas estatísticas para seleccionar a orde óptima do modelo ARIMA. A continuación, tomando como referencia Burnham e Anderson (2004), explicaremos un pouco máis cada un deles. A idea básica detrás de todos estes criterios é obter o máximo poder explicativo (medido polo valor da función de verosimilitud) con só uns poucos parámetros. Os criterios de selección do modelo escollen esencialmente o modelo co mellor axuste, medido pola función de verosimilitude, e está suxeito a un termo de penalización, para evitar un sobreaxuste, que aumenta co número de parámetros do modelo. Algunos dos criterios de información máis coñecidos son: Criterio de Información de Akaike (AIC), Criterio de Información de Akaike corrixido (AICc) e Criterio de información de Schwarz-Bayes (BIC).

As fórmulas para os criterios de selección de modelos empregadas por JDemetra+ son:

$$\begin{aligned} AIC &= -2L + 2n_p \\ AICc &= -2L + 2n_p + \frac{2n_p(n_p + 1)}{N - n_p - 1} \\ BIC &= -2L + n_p \log N \end{aligned}$$

onde:

- N é o número de observacións da serie de tempo.
- n_p é o número de parámetros estimados.
- L é a función de logverosimilitude.

Para cada criterio de selección de modelo prefírese o modelo que devolve o valor máis pequeno. Como o criterio AIC está sesgado para mostras pequenas a miúdo é substituído por AICc. Cómpre sinalar que o criterio BIC impón un termo de penalización maior que AIC, polo que BIC tende a seleccionar modelos más sinxelos que os elixidos por AIC. Ademais, a diferenza entre ambos criterios pode ser enorme se N é grande.

2.2. Estimación das componentes principais

A continuación, centrarémonos na segunda parte do axuste da serie de tempo, consistente en descompoñer a serie linealizada en componentes principais e estimálas. Aquí atopámonos coa principal diferenza entre os dous métodos implementados en JDemetra+, onde distinguiremos entre a metodoloxía SEATS e X-11, en función de se estamos a traballar con TRAMO-SEATS ou X-13ARIMA-SEATS, respectivamente.

2.2.1. SEATS

Comezaremos falando de SEATS, que é un programa para estimar compoñentes non observadas nunha serie temporal, a partir de modelos ARIMA. SEATS descompón a serie linealizada nas compoñentes de tendencia, estacional e irregular, acompañadas dunha compoñente de erro (ruído branco). Ademais, proporciona predicións para estas compoñentes, xunto cos errores estándar asociados e, por último, asigna os efectos deterministas a cada compoñente dando lugar as compoñentes finais. Os estimadores das compoñentes calcúlanse con filtros de Wiener-Kolmogorov, a partir da información proporcionada por TRAMO.

O programa SEATS parte da suposición de que a serie temporal linealizada, x_t , segue o modelo ARIMA,

$$\phi(B)\delta(B)x_t = \theta(B)a_t, \quad (2.5)$$

que se pode reescribir de forma más compacta denotando $\phi(B)\delta(B)$ por $\varphi(B)$,

$$\varphi(B)x_t = \theta(B)a_t, \quad (2.6)$$

é dicir, $\varphi(B)$, ao que nos referiremos como polinomio AR completo (abuso de notación), virá dado por:

$$\begin{aligned} \varphi(B) &= \phi(B)\delta(B) = \phi_p(B)\Phi_P(B^s)(1 - B)^d(1 - B^s)^D \\ &= (1 + \phi_1 B + \dots + \phi_p B^p)(1 + \Phi_1 B + \dots + \Phi_P B^{Ps})(1 - B)^d(1 - B^s)^D. \end{aligned}$$

En particular, como xa adiantabamos ao introducir os modelos ARIMA, en JDemetra+ tense que $p \leq 3$ e $P \leq 1$, polo que podemos reescribir $\varphi(B)$ de xeito más simple como

$$\varphi(B) = (1 + \phi_1 B + \phi_2 B^2 + \phi_3 B^3)(1 + \Phi_1 B^s)(1 - B)^d(1 - B^s)^D.$$

Nótese que a serie temporal estocástica x_t pódese predecir usando as súas observacións pasadas, cometendo un certo erro, dado por ruído branco, que é a parte que non se pode predecir en función da historia pasada da serie.

Agora consideremos o modelo de descomposición aditiva

$$x_t = \sum_{i=1}^k x_{it}, \quad (2.7)$$

onde $t = 1, \dots, T$, sendo T o número total de observacións da serie e o subíndice i refírese ás compoñentes ortogonais: tendencia, estacional, irregular e de erro, con k denotando o número total de ditas compoñentes. Ademais da compoñente de erro, que se supón como ruído branco, asúmese que cada compoñente segue o modelo ARIMA que se pode representar, usando notación análoga á da ecuación (2.6), como:

$$\varphi_i(B)x_{it} = \theta_i(B)a_{it}, \quad (2.8)$$

onde $\varphi_i(B) = \phi_i(B)\delta_i(B)$, x_{it} é a i -ésima compoñente non observada, $\varphi_i(B)$ e $\theta_i(B)$ son polinomios finitos de orde p_i e q_i , respectivamente, e a_{it} , a perturbación asociada con tal compoñente. Ademais, a_{it} e a_{jt} non están correlacionados para $i \neq j$ e para calquera t .

Logo, dado que a agregación de modelos ARIMA produce modelos ARIMA, a serie x_t tamén seguirá un modelo ARIMA. Polo tanto, a idea radica en considerar que cada compoñente se pode representar por un modelo ARIMA, sendo todos eles compatibles coa serie resultante da súa agregación, de tal forma que o modelo ARIMA identificado e estimado para a serie observada x_t descomporase dando paso a un modelo para cada compoñente. Podemos escribir formalmente este resultado tendo en conta que $x_t = \frac{\theta(B)}{\varphi(B)}a_t$ e $x_{it} = \frac{\theta_i(B)}{\varphi_i(B)}a_{it}$, e substituíndo na ecuación (2.7),

$$\frac{\theta(B)}{\varphi(B)}a_t = \sum_{i=1}^k \frac{\theta_i(B)}{\varphi_i(B)}a_{it}, \quad (2.9)$$

de onde podemos extraer as relacións

$$\varphi(B) = \prod_{i=1}^k \varphi_i(B),$$

$$\theta(B)a_t = \sum_{i=1}^k \prod_{j=1, j \neq i}^k \varphi_j(B)\theta_i(B)a_{it}.$$

Logo os polinomios AR para as componentes, $\varphi_i(B)$, derívanse facilmente mediante a factorización do polinomio AR $\varphi(B)$:

$$\varphi(B) = \prod_{i=1}^k \varphi_i(B), \quad (2.10)$$

mentres que os polinomios MA para as componentes, xunto coas varianzas das innovacións $V(a_i)$, non se poden obter simplemente mediante a relación:

$$\theta(B)a_t = \sum_{i=1}^k \varphi_{ni}(B)\theta_i(B)a_{it}, \quad (2.11)$$

onde $\varphi_{ni}(B)$ é o produto de todo $\varphi_j(B)$, $j = 1, \dots, k$, excepto de $\varphi_i(B)$. Neste caso, son necesarias outras suposicións para facer fronte ao problema de subidentificación, xa que, existe un número infinito de descomposicións admisibles. Polo tanto, os polinomios MA e as varianzas de innovación áinda non se poden identificar a partir do modelo de x_t . Para resolver este problema de subidentificación e identificar unha descomposición única, supонse que para cada componente a orde do polinomio MA non é maior que a orde do polinomio AR ($p_i \leq q_i$) e aplícase a solución canónica, é dicir, todo o ruído branco aditivo engádese á componente de erro.

Para comprender como SEATS factoriza os polinomios AR completos na ecuación (2.10), deterémonos na obtención das súas raíces.

Volvendo á ecuación inicial (2.6), multiplícanse por x_{t-k} , con $k > q$, ambos lados da igualdade, e calculamos as esperanzas,

$$\mathbb{E}[\varphi(B)x_t x_{t-k}] = \mathbb{E}[\theta(B)a_t x_{t-k}] \quad (2.12)$$

o lado derecho da ecuación desaparece:

$$\mathbb{E}[\theta(B)a_t x_{t-k}] = \theta(B)\mathbb{E}[a_t x_{t-k}] = \theta(B)\mathbb{E}(a_t)\mathbb{E}(x_{t-k}) = 0$$

por ser a_t ruído branco (con media cero) independente do proceso estocástico x_{t-k} . O lado esquerdo, se tomamos esperanzas, pasa a ser:

$$\mathbb{E}[\varphi(B)x_t x_{t-k}] = \varphi(B)\mathbb{E}[x_t x_{t-k}] = \varphi(B)\gamma_k$$

onde γ_k denota a función de autocovarianzas de x_{t-k} e x_t e B opera sobre o subíndice k . Substituíndo en (2.12), chegamos a

$$\gamma_k + \varphi_1\gamma_{k-1} + \dots + \varphi_p\gamma_{k-p} = 0. \quad (2.13)$$

A función γ_k é unha solución da ecuación diferencial homoxénea (2.13) con ecuación característica:

$$z^p + \varphi_1 z^{p-1} + \dots + \varphi_{p-1} z + \varphi_p = 0. \quad (2.14)$$

A ecuación característica ten raíces reais e complexas. Recordemos que un número complexo $x = a + bi$, con a e b números reais, pódese representar como $x = r(\cos(\omega) + i \sen(\omega))$, onde i é a unidade imaxinaria ($i^2 = -1$), r é o módulo de x , é dicir, $r = |x| = \sqrt{a^2 + b^2}$, e ω é o argumento ou frecuencia.

En particular, se denotamos z_1, \dots, z_p ás raíces da ecuación (2.14), estamos ante as inversas das raíces do polinomio $\varphi(B) = 0$, é dicir, $z_i = B_i^{-1}$.

Apoiándonos nas referencias Caporello et al. (2001) e Novás et al. (2021), tense que o polinomio completo AR, $\varphi(B)$, asigna as súas raíces ás componentes segundo o seu módulo e frecuencia asociados do seguinte xeito:

- As raíces de $(1 - B)^d$ son asignadas á compoñente de tendencia.
- As raíces de $(1 - B^s)^D = ((1 - B)(1 + B + \dots + B^{s-1}))^D$ asígnanse:
 1. Raíz de $(1 - B)^D$ á compoñente de tendencia.
 2. Raíces de $(1 + B + \dots + B^{s-1})^D$ á compoñente estacional.
- As raíces de $\phi_p(B)$ asígnanse en función do comportamento das raíces do polinomio

$$\phi_p(z) = z^p + \phi_1 z^{p-1} + \dots + \phi_p \quad \text{con } z = B^{-1}$$

do seguinte xeito:

1. Raíces reais positivas
 - a) Se módulo $\geq k_1$, asígnase á tendencia.
 - b) Se módulo $< k_1$, asígnase á compoñente irregular.

En concreto, en JDemetra+, k_1 por defecto vale 0.5.
2. Raíces reais negativas
 - a) Se $s \neq 1$, e módulo $\geq k_2$ asígnase á compoñente estacional.
 - b) Se $s \neq 1$ e módulo $< k_2$ asígnase á compoñente irregular.
 - c) Se $s = 1$ (datos anuais), asígnase á compoñente irregular.

En concreto, en JDemetra+, k_2 por defecto vale 0.8, ou 0.9 en caso de que $\phi_p(B)$ non conteña ningunha outra raíz estacional.
3. Raíces complexas
 - a) Se a frecuencia, w , se atopa comprendida nun intervalo próximo á frecuencia estacional asígnase á compoñente estacional.
 - b) Noutro caso asígnase á compoñente irregular.

- As raíces de $\Phi_P(B)$ asígnanse en función do comportamento das raíces de $\Phi_P(z) = z^s + \phi_s$ con $z = B^{-1}$.
 1. Se $\phi_s > 0$, as raíces asígnanse á compoñente irregular.
 2. Se $\phi_s < 0$, denotemos por ϕ a raíz real positiva de $(-\phi_s)^{1/s}$, tal que o polinomio $\Phi_P(z)$ pódese reescribir como

$$\Phi_P(z) = (z - \phi)(z^{s-1} + \phi z^{s-2} + \phi^2 z^{s-3} + \dots + \phi^{s-1}).$$

- a) Cando $D = 1$, a raíz AR $(1 - \alpha B)$ asígnase á tendencia e as outras $s - 1$ raíces á compoñente estacional.
- b) Cando $D = 0$, a raíz asígnase á compoñente estacional se $\phi_s < -0.2$ ou se os test de estacionabilidade indican a presenza desta. Noutro caso, a raíz asígnase á compoñente irregular.

En consecuencia, podemos apreciar que a factorización inicial do polinomio AR nos deu como resultado a identificación dos polinomios AR para as compoñentes principais de tendencia, estacionalidade e irregular, que conteñen as correspondentes raíces AR asociadas á cada unha, chegando á ecuación (2.10).

Por exemplo, se consideramos o modelo de liñas aéreas para unha serie temporal mensual:

$$(1 - B)(1 - B^{12})x_t = (1 + \theta_1 B)(1 + \Theta_1 B^{12})a_t,$$

modelo que se usa moi frecuentemente para modelar series de tempo económicas xa que é parsimonioso e relativamente fácil de estimar, tense

1. A raíz $(1 - B)$ asígnanse á compoñente de tendencia.
2. As raíces de $(1 - B^{12}) = (1 - B)(1 + B + \dots + B^{11})$ asignanse de xeito que $(1 - B)$ se asocia coa tendencia e $(1 + B + \dots + B^{11})$ coa compoñente estacional.
3. Non temos parte AR estacionaria nin estacional, polo que tampouco raíces asociadas.

Polo tanto a descomposición dada por SEATS no modelo para a compoñente de tendencia quedaría como

$$(1 - B)(1 - B)c_t = (1 + \theta_{c,1}B + \theta_{c,2}B^2)a_{c,t}$$

e o modelo para a compoñente estacional sería

$$(1 + B + \dots + B^{11})s_t = (1 + \theta_{s,1}B + \dots + \theta_{s,11}B^{11})a_{s,t}.$$

Como resultado, o modelo da liña aérea descompónse da seguinte forma:

$$\frac{(1 + \theta_1B)(1 + \Theta_1B^{12})}{(1 - B)(1 - B^{12})}a_t = \frac{(1 + \theta_{s,1}B + \dots + \theta_{s,11}B^{11})}{(1 + B + \dots + B^{11})}a_{s,t} + \frac{(1 + \theta_{c,1}B + \theta_{c,2}B^2)}{(1 - B)(1 - B)}a_{c,t} + u_t.$$

A compoñente irregular non está presente neste caso e a compoñente de erro é o ruído branco.

Unha vez temos definido un modelo teórico para as compoñentes, é preciso estimalas, isto é, obter as series temporais para cada x_{it} a partir dos datos observados de x_t . Este proceso realizaase mediante filtros de Wiener-Kolmogorov (WK). Levamos a cabo un filtrado de x_t da forma:

$$\hat{x}_{it} = v_i(B, F)x_t,$$

onde os filtros $v_i(B, F)$, con $F = B^{-1}$, empregados por SEATS, pertenecentes á familia Wiener-Kolmogorov, tratan de minimizar o erro cuadrático medio entre o estimador e a compoñente teórica. Formalmente, pódense representar como a solución dun problema de optimización restrinxida:

$$\begin{aligned} \min \quad & \mathbb{E}(x_{it} - \hat{x}_{it})^2 \\ \text{s.a.} \quad & x_{it} = \frac{\theta_i(B)}{\phi_i(B)}a_{it} \quad i = 1, \dots, k. \end{aligned}$$

A solución deste problema permitirános obter estimacións para cada x_{it} da forma:

$$\hat{x}_{it} = k_i \frac{\psi_i(B)\psi_i(F)}{\psi(B)\psi(F)}x_t,$$

onde $\psi(B) = \frac{\theta(B)}{\phi(B)}$ e $k_i = \frac{V(a_i)}{V(a)}$, sendo $V(a_i)$ e $V(a)$ as varianzas de a_{it} e a_t respectivamente.

A modo de exemplo, se consideramos un modelo de dúas compoñentes, tendo por un lado a serie desestacionalizada (x_{des}), que será o sinal de interese, e por outro a compoñente estacional (x_{est}) ou o resto (ruído). Expresando os polinomios $\psi(B)$ como funcións dos polinomios AR e MA, despois da cancelación de raíces, o estimador de x_{des} pódese representar como:

$$\hat{x}_{des} = k_{des} \frac{\theta_{des}(B)\theta_{des}(F)\varphi_{est}(B)\varphi_{est}(F)}{\theta(B)\theta(F)}x_t,$$

onde

$$v_{des}(B, F) = k_{des} \frac{\theta_{des}(B)\theta_{des}(F)\varphi_{est}(B)\varphi_{est}(F)}{\theta(B)\theta(F)}$$

é un filtro WK. Estes filtros son simétricos e centrados, e a invertibilidade de x_t garantiza a converxencia de ditos filtros. En particular, a derivación das compoñentes require dunha serie infinita x_t , non obstante, a converxencia do filtro garante que na práctica dita serie poida ser aproximada mediante un filtro finito (truncado).

Acabamos de apreciar que o filtro WK depende tanto da compoñente a estimar como dos modelos da serie. En consecuencia, o estimador da compoñente e o filtro WK reflecten as características dos datos e, por construcción, o filtro adáptase para cada serie considerada. Na práctica, dito filtro aplícase a x_t estendido con predicións e retrocesos do modelo ARIMA. Polo tanto, o modelo ARIMA é de especial importancia para o método SEATS e a súa especificación incorrecta dá lugar a unha descomposición incorrecta. Para unha explicación máis detallada do procedemento de descomposición das series de tempo pódese consultar Maravall (1987).

2.2.2. X-11

Unha vez visto o procedemento levado a cabo por SEATS, a continuación centrarémonos na descomposición da serie de tempo en compoñentes principais empregando a metodoloxía X-11.

O programa X-11 é o resultado dunha longa tradición de suavización non paramétrica baseada en medias móbiles, entendendo por medias móbiles como as medias ponderadas nun intervalo móvil. As medias móbiles teñen dous inconvenientes importantes: non son resistentes no sentido de que se poden ver profundamente afectados por valores atípicos e o suavizado dos extremos da serie non se pode facer salvo con medias móbiles asimétricas que introducen desfasaes e atrasos na detección de puntos de inflexión. Para superar estes problemas aplícase a metodoloxía vista para a primeira parte do axuste dunha serie de tempo na que se empregaba un modelo RegARIMA para eliminar os efectos deterministas da serie e axustábase á serie un modelo ARIMA que se empregaría agora para calcular predicións. Logo, unha vez feito isto estamos en condicións de aplicar o algoritmo X-11 para calcular as compoñentes de tendencia, estacionalidade e irregular.

Definiremos a media móvil de coeficientes $\{\theta_k\}$ como:

$$M(X_t) = \sum_{k=-e}^{+f} \theta_k X_{t+k}.$$

É dicir, a idea consiste en substituír o valor en tempo t da serie por unha media ponderada de e valores pasados, o valor actual e f valores futuros da serie. A cantidade $e + f + 1$ chámase orde da media móvil. En particular, cabe mencionar dous casos particulares:

- Se $e = f$, é dicir, se o número de puntos no pasado é o mesmo que o número de puntos no futuro, dise que a media móvil está centrada.
- Se $e = f$ e $\theta_{-k} = \theta_k$ para calquera k , dise que a media móvil M é simétrica. En concreto, unha das medias móbiles más sinxelas é a media móvil simétrica de orde $E = 2e + 1$ onde todos os pesos son iguais a $\frac{1}{E}$.

Xeralmente, cunha media móvil de orde $e + f + 1$ calculada para o instante t con e puntos no pasado e f puntos no futuro, será imposible suavizar os primeiros e valores e os últimos f valores da serie debido á falta de datos de entrada na fórmula da media móvil. É dicir, esta fórmula de media móvil funciona ben para todas as observacións de series temporais excepto para os primeiros valores e e os últimos valores f .

No método X-11, as medias móbiles simétricas xogan un papel importante xa que non introducen ningún cambio de fase na serie suavizada. Pero, para evitar perder información nos extremos das series, complétanse con medias móbiles asimétricas ou aplícanse á serie ampliada mediante predicións.

Para a estimación da compoñente estacional, X-13ARIMA-SEATS utiliza $D \times E$ medias móbiles obtidas a partir da composición dunha media móvil simple de orde D , cuxos coeficientes son todos iguais a $\frac{1}{D}$, e dunha media móvil simple de orde E , cuxos coeficientes son todos iguais a $\frac{1}{E}$. Isto significa, praticamente, aplicar sucesivamente á serie as dúas medias móbiles.

As medias móbiles compostas son amplamente utilizadas polo método X-11. Por exemplo, para unha estimación inicial da tendencia, o método X-11 utiliza unha media móvil 2×4 no caso dunha serie temporal trimestral e unha media móvil 2×12 para unha mensual. En particular, a media móvil 2×4 é unha media de orde 5 con coeficientes $\frac{1}{8}\{1, 2, 2, 2, 1\}$, tal que, a suma dos coeficientes é igual a 1 e é simétrica, polo que

conserva as tendencias lineais. De forma xeral, unha media móbil $D \times E$ é unha media móbil simétrica de orde $D + E - 1$. As medias móbiles compostas tamén se utilizan para extraer o compoñente estacional. As que se empregan na execución automática do método (sen intervención do usuario) son 3×3 , 3×5 e 3×9 .

Na estimación da tendencia tamén se utilizan as medias móbiles de Henderson. As medias de Henderson equivalen a medias móviles ponderadas, pero onde os pesos se obteñen como resultado de supoñer que a compoñente de tendencia segue una parábola de terceiro grao nun intervalo de tempo curto (entre un e dous anos, aproximadamente). Estes filtros elixíronse polas súas propiedades de suavizado. Os coeficientes dunha media móbil de Henderson de orde $2e + 1$ pódense calcular mediante a fórmula:

$$\theta_i = \frac{315 [(e-1)^2 - i^2] [e^2 - i^2] [(e+1)^2 - i^2] [3e^2 - 16 - 11i^2]}{8e(e^2 - 1)(4e^2 - 1)(4e^2 - 9)(4e^2 - 25)}.$$

Tomando como referencia Ladiray e Queeneville (2000), a idea do algoritmo de desestacionalización que segue X-11 pódese plantexar en catro etapas. En particular, para presentar a súa estrutura, denotaremos por M_0 e M_1 ás medias móbiles empregadas para estimar a tendencia e a compoñente estacional, respectivamente. Ademais, utilizaremos o superíndice (i) para indicar que estamos ante a estimación i -ésima da compoñente.

Paso 1 Estimación da tendencia por media móbil:

$$T_t^{(1)} = M_0(X_t).$$

Paso 2 Estimación do compoñente estacional-irregular:

$$(S_t + I_t)^{(1)} = X_t - T_t^{(1)}.$$

Paso 3 Estimación do compoñente estacional mediante media móbil sobre cada mes:

$$S_t^{(1)} = M_1 \left[(S_t + I_t)^{(1)} \right],$$

e en consecuencia

$$I_t^{(1)} = (S_t + I_t)^{(1)} - S_t^{(1)}.$$

Nesta etapa trátase de suavizar os valores da compoñente estacional-irregular de cada mes para extraer a evolución do coeficiente estacional do mes. A media móbil empregada terá que reproducir o mellor posible a compoñente estacional de cada mes, reducindo ao máximo a compoñente irregular.

Paso 4 Estimación da serie desestacionalizada:

$$SA_t^{(1)} = (T_t + I_t)^{(1)} = X_t - S_t^{(1)}.$$

Polo tanto, a única dificultade neste plantexamento reside en seleccionar adecuadamente as medias móbiles empregadas nos pasos 1 e 3. O método X-11 levará a cabo este algoritmo simple, utilizando medias móbiles cuidadosamente elixidas e afinando, pouco a pouco, as estimacións das compoñentes a través das iteracións do algoritmo. Entón, podemos definir o algoritmo do método X-11 dicindo que se corresponde co dobre uso consecutivo do algoritmo que acabamos de describir, cambiando de cada vez as medias móbiles utilizadas.

Por exemplo, os pasos a seguir para unha serie temporal mensual serían os seguintes:

Paso 1 Estimación da tendencia por media móbil 2×12 :

$$T_t^{(1)} = M_{2 \times 12}(X_t).$$

A media móbil utilizada nesta etapa é unha media móbil de 13 termos, chamada 2×12 , de coeficientes $\frac{1}{24} \{1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1\}$, que conserva a tendencia lineal, elimina as estacionalidades constantes de orde 12 e minimiza a varianza da parte irregular.

Paso 2 Estimación do compoñente estacional-irregular:

$$(S_t + I_t)^{(1)} = X_t - T_t^{(1)}.$$

Paso 3 Estimación do compoñente estacional mediante a media móbil 3×3 sobre cada mes:

$$S_t^{(1)} = M_{3 \times 3} [(S_t + I_t)^{(1)}].$$

A media móbil usada aquí é unha media móbil 3×3 en 5 termos, con coeficientes $\frac{1}{9}\{1, 2, 3, 2, 1\}$. A compoñente estacional céntrase despois cunha media móbil de 2×12 .

$$\tilde{S}_t^{(1)} = S_t^{(1)} - M_{2 \times 12} (S_t^{(1)}).$$

Paso 4 Estimación da serie desestacionalizada:

$$SA_t^{(1)} = (T_t + I_t)^{(1)} = X_t - \tilde{S}_t^{(1)}.$$

Esta primeira estimación da serie desestacionalizada debe, por construción, conter menos estacionalidade. O método X-11 executa de novo o algoritmo cambiando as medias móbiles para ter en conta esta propiedade.

Paso 5 Estimación da tendencia pola media móbil de Henderson de 13 termos:

$$T_t^{(2)} = H_{13} (SA_t^{(1)}).$$

As medias móbiles de Henderson, aínda que non teñen propiedades especiais en canto á eliminación da estacionalidade (a cal, neste punto, xa é praticamente inexistente ou moi reducida), teñen un moi bo poder de suavización e manteñen unha tendencia localmente polinómica de segundo grao.

Paso 6 Estimación da compoñente estacional-irregular:

$$(S_t + I_t)^{(2)} = X_t - T_t^{(2)}.$$

Paso 7 Estimación do compoñente estacional mediante unha media móbil 3×5 ao longo de cada mes:

$$S_t^{(2)} = M_{3 \times 5} [(S_t + I_t)^{(2)}].$$

A media móbil usada aquí é unha media móbil 3×5 en 7 termos, de coeficientes $\frac{1}{15}\{1, 2, 3, 3, 3, 2, 1\}$ e mantén tendencias lineais. A compoñente estacional céntrase despois cunha media móbil de 2×12 .

$$\tilde{S}_t^{(2)} = S_t^{(2)} - M_{2 \times 12} (S_t^{(2)}).$$

Paso 8 Estimación da serie desestacionalizada:

$$SA_t^{(2)} = (T_t + I_t)^{(2)} = X_t - \tilde{S}_t^{(2)}.$$

Toda a dificultade reside, pois, na elección das medias móbiles utilizadas para a estimación da tendencia nos pasos 1 e 5, por unha banda, e para a estimación da compoñente estacional nos pasos 3 e 7.

Ademais, para avaliar as distintas compoñentes dunha serie, tendo en conta a posible presenza de observacións extremas ou atípicas, X-11 procederá de forma iterativa: estima as compoñentes, busca na compoñente irregular os efectos que están perturbando a análise da serie, estima as compoñentes sobre a serie corrixida, busca os efectos que perturban a nova serie na compoñente irregular, e así sucesivamente ata que desaparezan ditos efectos por completo.

2.3. Comparación TRAMO-SEATS e X-13ARIMA-SEATS

Como ben se explica en Villarreal (2005), a día de hoxe non existe un criterio que permita escoller un método entre TRAMO-SEATS e X-13ARIMA-SEATS como o mellor para o axuste estacional, xa que ambos enfoques presentas vantaxes e desvantaxes. Por exemplo, en X-13ARIMA-SEATS os procedementos non se basean en modelos explícitos e non son óptimos, o que dificulta a inferencia estatística dos estimadores obtidos. En contraposición, en TRAMO-SEATS os procedementos son óptimos e permiten facer inferencia sobre os estimadores, pero os modelos teóricos para unha serie temporal non son realmente modelos económico, polo que existen dúbidas sobre a idoneidade á hora de modelar a serie con ditos modelos.

É importante mencionar que a teoría económica non ofrece unha descripción formal das compoñentes en que se desglosa unha serie temporal, o que fai praticamente imposible chegar a unha conclusión sobre a superioridade dun ou doutro método de axuste estacional, xa que, sen a descripción das compoñentes o problema de axuste estacional non é un problema ben definido.

Existen varios estudos nos que, baixo diversos criterios, tanto teóricos como empíricos, avaliouuse a rea-lización de ambos métodos. Probablemente o estudio comparativo máis completo é o realizado polo Eurostat (Planas, 1997). Dito estudo centrouse na análise dalgúns propiedades teóricas de ambos métodos sobre series de tempo económicas, co fin de destacar vantaxes e inconvenientes.

Axuste previo: A partir da comparación dos procedementos usados en RegARIMA e TRAMO conclúese que ambos métodos están baseados na mesma metodoloxía e que baixo supostos equivalentes se obteñen resultados similares. Con todo, cabe mencionar que no caso de que o coeficiente do polinomio AR estacional ($\Phi_P(\cdot)$) se atope cerca de 1, os resultados obtidos por TRAMO en mostras grandes teñen un menor erro cuadrático medio. Así mesmo, demostrouse que ademais de ser más rápida, a rutina automática usada en TRAMO é preferible á rutina semiautomática de RegARIMA.

Axuste estacional de series con comportamento extremo: En concreto, investigouse o axuste estacional cando o comportamento das compoñentes é case determinista e cando é moi inestable (ou non determinista).

Cando o comportamento da compoñente estacional é moi estable ou determinista, as raíces do polinomio MA ($\theta(\cdot)$) do modelo identificado son raíces unitarias, o que significa que o filtro WK pasa a ser indeterminado.

Ao comparar os resultados do procedemento utilizado para estimar modelos non invertibles en SEATS cos resultados obtidos mediante X-12-ARIMA, que pola súa natureza empírica nunca se confronta ao problema da invertibilidade, as conclusións foron que a rutina utilizada en SEATS é válida e ademais produce resultados más estables que os obtidos en X-12-ARIMA. Este resultado é atribuíble á relativa rixidez dos filtros dispoñibles no módulo X-11, que non permiten distribuír a variabilidade entre a compoñente estacional e a compoñente irregular de forma satisfactoria.

Por outra parte, cando os patróns das componentes non observadas son inestables ou non deterministas, o modelo identificado por TRAMO non admite unha descomposición aceptable, polo que ten que ser aproximado por SEATS. Igual que no caso anterior, X-12-ARIMA non se confronta a este problema. Comparáronse a aproximación utilizada por SEATS cos resultados obtidos por X-12-ARIMA, concluíndo que a pesar de que a aproximación utilizada por SEATS introduce variabilidade adicional á estimación das compoñentes, os resultados obtidos son más estables que os obtidos polo módulo X-11, o cal non é capaz de adaptarse á variabilidade da compoñente estacional.

Revisións dos datos axustados: Realizouse un exame sobre as revisións que sofren as compoñentes estimadas cando se agregan observacións á serie de tempo observada. Para o caso da compoñente estacional púxose de manifesto que SEATS en xeral resulta nunha compoñente más estable.

Aínda que se pode ver que a velocidade de converxencia en X-12-ARIMA é maior á velocidad observada en SEATS, atribúese este resultado ao tamaño das revisións. Con respecto á compoñente de tendencia, en xeral o espectro da compoñente de tendencia en SEATS tende a cero más rápido que en X-12-ARIMA.

Polo que a interpretación dos resultados non é tan clara, xa que as diferenzas débense a que as compoñentes miden diferentes aspectos da serie.

Estabilidade das compoñentes estimadas: Para avaliar este aspecto, desenvolveuse unha medida da distancia entre filtros, co fin de aproximar os filtros utilizados polo método X-11 co modelo de liñas aéreas. Unha vez que encontran unha correspondencia entre os filtros, realizan unha comparación do axuste estacional obtido con ambos métodos utilizando series que estaban ben descritas por este modelo.

Os resultados de dito traballo indican que tanto as series axustadas estacionalmente como as series de tendencia obtidas utilizando SEATS, son máis estables que as obtidas utilizando o módulo X-11. Este resultado chama a atención, xa que no proceso de identificación SEATS maximiza a variabilidade da compoñente irregular. Un estudo sobre a orixe da maior variabilidade obtida por X-11 con respecto á obtida por SEATS conclúe que se debe a que na maioría dos casos, a ganancia do filtro é maior a un entre as frecuencias estacionais e, dado que estas frecuencias están asociadas a movementos de curto prazo, tense como resultado que a serie axustada estacionalmente mostra máis variabilidade que a serie orixinal.

O Eurostat, extraendo as ideas principais deste estudo, que aparecen resumidas no Cadro 2.1, adoptou e recomenda TRAMO-SEATS para realizar o axuste estacional de series económicas, e en consecuencia dita metodoloxía é a que se está a empregar por norma xeral en Europa.

Sen embargo, cabe destacar que existen outros estudos nos que as conclusións non foron tan claras, polo que non é recomendable menosprezar a diversidade de ferramentas empíricas asociadas á familia X-11, as cales seguen sendo utilizadas pola maioría de Bancos Centrais e Oficinas de Estatística ao redor do mundo.

	TRAMO-SEATS	X-13ARIMA-SEATS
Axuste previo	✓	✓
Axuste estacional con compoñentes deterministas	✓	✗
Axuste estacional con compoñentes non deterministas	✗	✓
Revisións dos datos axustados	Non comparable	Non comparable
Estabilidade das compoñentes estimadas	✓	✗

Cadro 2.1: Comparativa TRAMO-SEATS e X-13ARIMA-SEATS.

2.4. Diagnose do modelo

Unha vez presentados os dous procedementos implementados en JDemetra+ para a corrección da estacionalidade dunha serie de tempo compre introducir unha metodoloxía que nos permita comprobar que o axuste realizado é adecuado aos datos da serie bruta dados. Para iso, o software JDemetra+ inclúe unha serie de test de diagnose, entre os cales enunciaremos e explicaremos brevemente aqueles que se mostran na versión de JDemetra+ adaptada ao entorno R Project (RJDemetra+), que en esencia, son os más relevantes.

Funcións de autocorrelación

Podemos comenzar por un análise gráfico das autocorrelacións dos residuos do modelo axustado. O conxunto de coeficientes de autocorrelación e de autocorrelación parcial, dispostos en función do retardo dan lugar á función de autocorrelación (ACF) e á función de autocorrelación parcial (PACF), respectivamente. Ditas

funcións de autocorrelación constitúen unha ferramenta valiosa para investigar as propiedades dunha serie temporal empírica. En JDemetra+ represéntanse os gráficos de cada unha das acompañadas dos seus intervalos de confianza, de modo que se o coeficiente de autocorrelación (ou autocorrelación parcial) está no intervalo de confianza considérase non significativo, en caso contrario, podería ser unha sinal da existencia de problemas no axuste da serie.

Normalidade dos residuos

Para avaliar a normalidade dos residuos, JDemetra+ emprega contrastes de proximidade da media e dos coeficientes de asimetria e curtose dos residuos da serie axustada aos valores correspondentes de ditos momentos para a distribución normal estándar (que valen 0, 0 e 3 respectivamente).

$$\begin{aligned} H_0 : \mu = 0 & \quad H_0 : \text{Asimetria} = 0 & H_0 : \text{Curtose} = 3 \\ H_1 : \mu \neq 0 & \quad H_1 : \text{Asimetria} \neq 0 & H_1 : \text{Curtose} \neq 3 \end{aligned}$$

Un valor significativo de calquera deles indicará que os residuos estandarizados non seguen unha distribución normal estándar. A normalidade dos residuos é crucial para a validez dos intervalos de predición. Polo tanto, o noso obxectivo será que os tres test saian favorables a H_0 , é dicir, non aceptar H_1 .

Independencia dos residuos

$$\begin{aligned} H_0 : \text{Os residuos do modelo son independentes,} \\ H_1 : \text{Os residuos do modelo non son independentes.} \end{aligned}$$

A independencia dos residuos avalíase a través do estadístico Q de Ljung-Box (Ljung e Box, 1978) calculado tanto para os retardos regulares como estacionais. Este contraste chequea a presenza de autocorrelación entre retardos. A presenza de autocorrelación nos residuos é sinal de que os valores de ditos residuos non son independentes. Polo tanto, neste contraste buscamos probar que os residuos son independentes, ou o que é o mesmo, non aceptar H_1 .

O número de retardos que se teñan en conta para o cálculo do estadístico considera os 24 primeiros retardos se a serie é mensual e os 16 primeiros se é trimestral. Para os retardos estacionais o contraste toma en ambos casos os dous primeiros retardos (retardo 12 e 24 para series mensuais e, retardo 4 e 8 para trimestrais).

O Q -estadístico de Ljung-Box vén dado por:

$$Q = N(N + 2) \sum_{k=1}^K \frac{\hat{\rho}_k^2}{N - k},$$

onde:

- $\hat{\rho}_k$ é o coeficiente de autocorrelación simple no retardo k dos residuos \hat{a}_t .
- N é o tamaño da mostra.
- K é o retardo máximo que se está a considerar, establecido en JDemetra+ en 24 (series mensuais) ou 8 (series trimestrais).

Se os residuos son aleatorios (que é o caso dos residuos dun modelo ben especificado), o estadístico Q distriuirase segundo unha χ^2_{K-m} , onde m é o número de parámetros no modelo que se axustou aos datos.

As probas Ljung-Box ás veces non rexeitan un modelo mal axustado. Polo tanto, hai que ter coidado de non aceptar un modelo en función dos seus resultados.

Linearidade dos residuos

$$\begin{aligned} H_0 &: \text{Os residuos do modelo son lineais,} \\ H_1 &: \text{Os residuos do modelo non son lineais.} \end{aligned}$$

Este test contrasta a presenza de autocorrelación na serie de cadrados dos residuos. A existencia de dita correlación implicaría que os residuos conteñen certa estrutura non lineal. Este contraste lévase a cabo co estadístico Q de Ljung-Box dos residuos ao cadrado (\hat{a}_t^2), tamén para os 24 ou 16 primeiros retardos dependendo de se a serie é mensual ou trimestral. Neste contraste buscamos probar que os residuos son independentes, ou o que é o mesmo, non aceptar H_1 .

Test de Kruskal-Wallis

$$\begin{aligned} H_0 &: \text{As medias son iguais en todos os períodos (serie sen estacionalidade),} \\ H_1 &: \text{As medias non son iguais en todos os períodos (serie con estacionalidade).} \end{aligned}$$

A proba de Kruskal-Wallis (Kruskal e Wallis, 1952) é unha proba non paramétrica utilizada para comprobar se as mostras se orixinan da mesma distribución. No noso caso, a hipótese nula afirma que todos os meses (ou trimestres) teñen a mesma media. Cando se rexeita esta hipótese, asúmese que os valores das series temporais difiren significativamente entre períodos.

A proba aplícase normalmente a k grupos de datos $\{x_i\}_j$ onde cada grupo $j = 1, \dots, k$ está composto por n_j observacións, que están indexadas por $i = 1, \dots, n_j$. Cada mes (ou trimestre) agrupa todas as observacións disponíbeis para un determinado número de anos.

O estadístico da proba vén dado por

$$Q = \frac{SS_t}{SS_e}$$

onde $SS_t = (N - 1) \sum_{j=1}^k n_j (\bar{r}_{.j} - \bar{r})^2$ e $SS_e = \sum_{j=1}^k \sum_{i=1}^{n_j} n_i (\bar{r}_{ij} - \bar{r})^2$

- $N = n_1 + \dots + n_k$ é o número total de observacións entre todos os grupos.
- n_j é o número de observacións no grupo j .
- r_{ij} é o rango da observación i no grupo j .
- $\bar{r}_{.j}$ é a media dos rangos absolutos dos datos do grupo j .
- O rango medio é $\bar{r} = \frac{1}{2}(N + 1)$.

Baixo a hipótese nula de que todos os grupos se xeran a partir da mesma distribución, o estadístico de proba Q é aproximado mediante unha distribución χ^2_{k-1} , onde k é o número de grupos a comparar. Se o número de grupos é pequeno ($k < 3$) entón esta proba pode ser engañosa polo que teríamos que pasar a utilizar a proba de Mann-Whitney-Wilcoxon (Neuhäuser, 2011), que non é máis que a versión non paramétrica da proba t de Student, aplicada a dúas mostras independentes. Se a proba é significativa hai evidencias de diferencias entre períodos.

Neste caso, como polo xeral estamos a traballar con series que presentan estacionalidade o resultado buscado será rexeitar H_0 , é dicir, hai probas significativas a favor da presenza de estacionalidade na serie.

Test para a presenza de estacionalidade asumindo estabilidade

H_0 : As medianas son iguais en todos os períodos (serie sen estacionalidade estable),

H_1 : As medianas non son iguais en todos os períodos (serie con estacionalidade estable).

Este test emprega o mesmo estatístico e as mesmas hipóteses de contraste que o test de Friedman (Friedman, 1937), polo que a continuación introducimos a información de dito test.

O test de Friedman é un método non paramétrico para comprobar que as mostras son extraídas da mesma poboación ou de poboacións con medianas iguais, é dicir, compróbase a importancia do efecto mes (ou trimestre).

Sexa a matriz de datos $\{x_{ij}\}_{n \times k}$ con n filas (o número de anos na mostra), k columnas (12 meses ou 4 trimestres, dependendo da frecuencia dos datos). Dita matriz debe ser substituída por unha nova matriz $\{r_{ij}\}_{n \times k}$, onde a entrada r_{ij} é o rango de x_{ij} dentro do bloque i .

O estatístico do test vén dado por

$$Q = \frac{SS_t}{SS_e}$$

onde $SS_t = n \sum_{j=1}^k (\bar{r}_{\cdot j} - \bar{r})^2$ e $SS_e = \frac{1}{n(k-1)} \sum_{i=1}^n \sum_{j=1}^k (r_{ij} - \bar{r})^2$.

- $\bar{r}_{\cdot j}$ son os rangos medios de cada mes (ou trimestre) j dentro de cada ano.

- O rango medio vén dado por $\bar{r} = \frac{1}{nk} \sum_{i=1}^n \sum_{j=1}^k r_{ij}$.

Para n ou k grande, é dicir, $n > 15$ ou $k > 4$, a distribución de probabilidade de Q pódese aproximar mediante a dunha distribución χ_{k-1}^2 .

Ao igual que pasaba co test de Kruskal-Wallis, como estamos a traballar con series que presentan estacionalidade o resultado buscado será rexeitar H_0 , é dicir, hai probas significativas a favor da presenza de estacionalidade estable na serie.

Test de estacionalidade evolutiva (estacionalidade en movemento)

H_0 : Non hai cambios na estacionalidade ao longo dos anos (serie estacional),

H_1 : Hai cambios na estacionalidade ao longo dos anos (serie non estacional).

A proba de estacionalidade evolutiva baséase nun modelo de Análise da Varianza Bidireccional (Fujikoshi, 1993), que se emprega para determinar se hai ou non unha diferenza estatisticamente significativa entre as medias de tres ou máis grupos independentes que se dividiron en dous factores. O modelo utiliza só os valores de anos completos e a hipótese nula valida que non haxa cambios na estacionalidade ao longo dos anos.

Sexa a matriz de datos $\{x_{ij}\}_{n \times k}$ con n filas (o número de anos na mostra), k columnas (12 meses ou 4 trimestres, dependendo da frecuencia dos datos).

Denotaremos cada observación por x_{ij} onde $j = 1, \dots, k$ denota o período (polo tanto $k = 12$ para series mensuais e $k = 4$ para series trimestrais), e o índice i con $i = 1, \dots, n$ denota o ano, tal que, n é o número de anos (completos) considerados.

A hipótese nula chequese co seguinte estatístico de test:

$$F_M = \frac{\frac{S_A^2}{n-1}}{\frac{S_R^2}{N-n-k+1}},$$

onde $S_A^2 = k \sum_{i=1}^n (\bar{x}_{i\bullet} - \bar{x}_{\bullet\bullet})^2$ e $S_R^2 = \sum_{i=1}^n \sum_{j=1}^k (x_{ij} - \bar{x}_{(ij)})^2$, sendo

- $\bar{x}_{\bullet\bullet}$ a media global.
- $\bar{x}_{i\bullet}$ a media mostral das observacións procedentes do ano i .
- $\bar{x}_{(ij)}$ a media mostral das observacións sen considerar o dato x_{ij} .
- N o tamaño da mostra.
- n número de anos completos.
- k número de períodos.

O estatístico F_M segue unha distribución $F_{n-1, N-n-k+1}$ con $(n-1)$ e $(N-n-k+1)$ graos de liberdade.

Segundo a liña dos dous test presentados anteriormente, como estamos a traballar con series que presentan estacionalidade o resultado buscado será non aceptar H_1 , é dicir, non hai probas significativas a favor de que a serie non é estacional.

Tests de estacionalidade residual

Tamén temos un conxunto de contrastes para chequear a presenza de estacionalidade residual na serie axustada estacionalmente (á que nos referiremos como “on sa”) e na compoñente irregular de dita serie (á que nos referiremos como “on i”). Por defecto, empréganse os test F e QS que se explican a continuación. Logo as hipóteses son:

$$\begin{aligned} H_0 &: \text{Non hai estacionalidade residual,} \\ H_1 &: \text{Hai estacionalidade residual.} \end{aligned}$$

A hipótese nula é a non presenza de estacionalidade residual. Ademais, móstrase o resultado dos F-test empregado para chequear a presenza de efectos residuais de trading day para a serie axustada estacionalmente (“on sa (td)”) e para a compoñente irregular (“on i (td)”). En concreto, en R, denótanse aos F-test e QS-test por f-test e qs-test, respectivamente, e imprímense os resultados das seguintes probas: qs test on sa, qs test on i, f-test on sa (seasonal dummies), f-test on i (seasonal dummies), Residual seasonality (entire series), Residual seasonality (last 3 years), f-test on sa (td) e f-test on i (td).

QS-test

$$\begin{aligned} H_0 &: \text{Residuos distribuídos de forma independente (non estacionais),} \\ H_1 &: \text{Residuos non distribuídos de forma independente (estacionais).} \end{aligned}$$

O test QS (Maravall, 2012) é unha variante da proba Ljung-Box (Ljung e Box, 1978) calculada en retardos estacionais, onde só consideramos autocorrelacións positivas. Máis exactamente,

$$QS = N(N+2) \sum_{k=1}^K \frac{[\max(0, \hat{\gamma}_{k,l})]^2}{n - k \cdot l},$$

onde $K = 2$, polo que só se consideran os primeiros e segundos retardos estacionais. Así, a proba comprobaría a correlación entre a observación real e as observacións atrasadas un e dous anos. Teñamos en conta que $l = 12$ cando se trata de observacións mensuais, polo que consideramos só as autocovarianzas $\hat{\gamma}_{12}$ e $\hat{\gamma}_{24}$. Pola súa banda, $l = 4$ no caso dos datos trimestrais.

Baixo H_0 , que indica que os datos están distribuídos de forma independente, o estatístico segue unha distribución χ^2_K , é dicir, χ^2_2 .

F-test

H_0 : As medias son iguais en todos os períodos (residuos non estacionais),

H_1 : As medias non son iguais en todos os períodos (residuos estacionais).

Por último, presentaremos o test F, que se basea no modelo ANOVA (Tejedor, 1999). A hipótese nula de dito test postula que as medias de todos os grupos a comparar son iguais. No noso caso permítanos estudar se existe diferenza entre grupos (meses ou trimestres). O estatístico do test vén dado por:

$$F = \frac{\frac{RSS_1}{k-1}}{\frac{RSS}{N-k}}$$

onde $RSS_1 = \sum_{j=1}^k n_j (\bar{res}_{\bullet j} - \bar{res}_{\bullet\bullet})^2$ e $RSS = \sum_{j=1}^k \sum_{i=1}^{n_j} (res_{ij} - \bar{res}_{\bullet j})^2$, sendo

- res_{ij} denota o residuo pertencente ao grupo j con $j = 1, \dots, k$ e na posición i , que indica o número de elementos que hai en cada grupo de residuos, con $i = 1, \dots, n_j$.
- $\bar{res}_{\bullet\bullet}$ a media global.
- $\bar{res}_{\bullet j}$ a media mostral das observacións asociadas co grupo j .

Este estatístico segue unha distribución F con $k - 1$ e $N - k$ graos de liberdade, sendo k o número de grupos (12 para o caso mensual e 4 para o trimestral) e $N = \sum_{j=1}^k n_j$ o tamaño da mostra.

Capítulo 3

Resultados

Unha vez establecidos os procedementos teóricos a seguir, o seguinte paso será levalos á práctica, é dicir, levar a cabo a modelización das series de tempo reais. Neste capítulo trataremos de explicar o proceso levado a cabo para conseguir o axuste máis adecuado en cada caso, que nos permitirá obter unha serie máis simple e interpretable.

Como punto de partida á hora de modelizar cada serie de tempo considéranse os parámetros preestablecidos por defecto en JDemetra+, engadíndolle o calendario de festividades de Galicia. Este paso é fundamental xa que o calendario que trae JDemetra+ por defecto, como xa comentamos, só recolle como festivos os días da fin de semana, sábados e domingos. En particular, para crear o calendario de Galicia consideraremos os seguintes días festivos (que terán unha consideración igual a un domingo):

- Aninovo, 1 de xaneiro.
- Día de Reis, 6 de xaneiro.
- Día internacional dos traballadores, 1 de maio.
- Día das letras galegas, 17 de maio.
- Día nacional de Galicia, 25 de xullo.
- Asunción, 15 de agosto.
- Festa nacional de España, 12 de outubro.
- Día de todos os santos, 1 de novembro.
- Día da Constitución, 6 de decembro.
- Día da Inmaculada Concepción, 8 de decembro.
- Nadal, 25 de decembro.

A maiores temos dous festividades que non se computan como festivos todos os anos, polo tanto, estas datas temos que definilas en períodos reducidos de 1 ano, é dicir, manualmente temos que incluír en que ano se contabilizan como festivos e en cales non (que é a consideración por defecto). Nesta situación temos:

- Día de San Xosé, 19 de marzo.
- Día de San Xoán, 24 de xuño.

Entre as especificacións por defecto de JDemetra+ as máis significativas son que realiza un test para escollear se aplica logaritmos ou non sobre a serie, que calcula os valores atípicos e o modelo ARIMA de forma automática e que a nivel calendario realiza unha proba para decidir se considerará Working Days, é dicir, só constrúe un regresor, contendo o efecto da diferenza entre os días laborables e non laborables, ou Trading Days (seis regresores). Con todo, cabe destacar que tende a considerar Working Days xa que Trading Days implica engadir complexidade ao modelo. Por defecto tamén se realiza un test para decidir se se inclúe no modelo o efecto de Working Days (Trading Days) ou non, e outro sobre a utilización do regresor de efecto de Semana Santa.

A idea inicial sería executar a serie bruta de datos con estes parámetros preestablecidos e analizar os resultados. En función do que obtemos podemos quedar co modelo obtido por defecto ou comenzar a facer modificacións sobre este, polo xeral modificar os outliers ou o modelo ARIMA pódennos resolver certos problemas que apareceran co modelo por defecto. Ademais, todas as series modelizadas neste traballo empregan a metodoloxía TRAMO-SEATS, xa que, aparte de ser a recomendada por Eurostat, X-13ARIMA-SEATS non aporta melloras significativas.

Para ver o procedemento a seguir de forma máis detallada, a continuación centrarémonos na análise de cinco series das incluídas entre as más relevantes para explicar a conjuntura económica de Galicia:

- Poboación ocupada en Galicia (EPA),
- Índice de producción industrial (IPI),
- Transporte marítimo,
- Ocupación nos establecementos hoteleiros (entradas),
- Xeración de enerxía eléctrica.

Ditas series foron escollidas porque abranguen diferentes ámbitos da conjuntura económica de Galicia ao mesmo tempo que poñen de manifesto as diferentes situacions ou casuísticas que se poden dar no proceso de axuste estacional. Toda a información de interese sobre ditas series foi recollida do Instituto Galego de Estatística (2022) e a metodoloxía empregada para levar a cabo o axuste estacional segue as indicacións do Instituto Nacional de Estadística (2019). Ademais, para traballar con JDemetra+ consultouse o manual do Departamento de Metodología y Desarrollo de la Producción Estadística (2019).

3.1. Poboación ocupada en Galicia (EPA)

O primeiro caso que consideraremos será o da serie de ocupados da Enquisa de poboación activa (EPA11_OCUT), que ten como fonte de referencia a Enquisa de poboación activa do IGE-INE.

Defínese como ocupados a aquelas persoas de 16 anos ou máis que durante a semana de referencia (preestablecida na enquisa) tiveron un traballo por conta allea ou exerceron unha actividade por conta propia, nalgúnha das seguintes situacións:

- Traballando polo menos unha hora por un soldo, salario ou beneficio empresarial ou ganancia familiar en metálico ou en especie.
- Con emprego pero sen traballar por razóns de enfermidade ou accidente, vacacións, festas, mal tempo ou razóns análogas.

Nótese que esta serie é trimestral, é dicir, temos datos por grupos de tres meses, polo que hai catro observacións por ano, dispoñendo de datos dende o primeiro trimestre do 1996 ata o primeiro trimestre do 2022. Representamos a serie dos datos brutos na Figura 3.1.

Axuste automático

Ao executar o axuste da serie de tempo cos parámetros preestablecidos, na pestaña de resultados principais o primeiro que obtemos é a seguinte advertencia: “decomposition. Model decomposition: Parameters cut

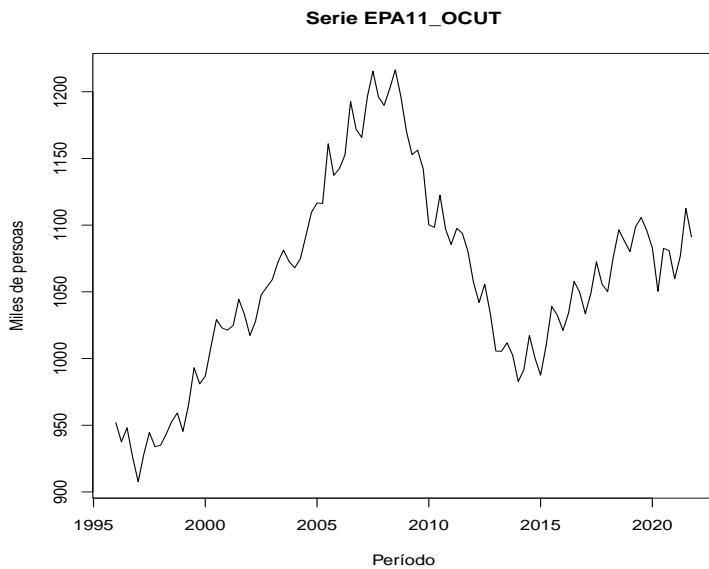


Figura 3.1: Serie orixinal de EPA11_OCUT.

off”. Nunca se poderá deixar un modelo con este erro presente, xa que nos está indicando un problema no módulo dalgunha raíz da parte MA, provocando que o modelo ARIMA seleccionado por SEATS non sexa o mesmo que o seleccionado por TRAMO, senón que os coeficientes da parte MA foron truncados. En consecuencia, xa sabemos que sempre que obteñamos esta advertencia non será aceptable o modelo, teremos que facerlle modificacións ata conseguir correxir dito problema.

Análise gráfica e diagnose do modelo

Podemos observar que os resultados de validación do modelo son favorables polo que non deberíamos alonxarnos demasiado do modelo que foi definido automaticamente. É dicir, o modelo calculado devolve bons resultados en canto a diagnose e axusta ben a serie gráficamente pero ten un problema de descomposición que debemos correxir. Unha idea para correxilo pode ser modificar o modelo ARIMA ou os atípicos obtidos automaticamente. O modelo obtido é un ARIMA(3,1,0)(0,1,1) sen media e os atípicos considerados son: un Atípico Aditivo (AO) no cuarto trimestre de 2005 e un Cambio Transitorio (TC) no segundo trimestre de 2020.

Graficamente, na Figura 3.2, podemos apreciar que a serie da tendencia no 2020 non se axusta todo o ben que debería ao comportamento real da serie, polo que a nosa primeira idea podería ser modificar os outliers establecidos por defecto.

Novo modelo construído a partir da intervención do usuario

Comezaremos considerando un caso moi simple dado por un único outlier asociado coa aparición do COVID-19. Como este cambio modifica a serie de forma permanente, consideraremos un Cambio de Nivel (LS) no segundo trimestre de 2020, asociado á aparición do COVID-19. Á hora de definir os atípicos manualmente en Pre-specified outliers temos un calendario que nos permite indicar a data exacta, así como, o tipo de atípico que queremos considerar (incluso se poden considerar varios tipos de outliers no mesmo día, pero non tería moito sentido). En particular, este calendario presenta unha opción de frecuencia, na que no noso caso debemos escoller frecuencia igual a catro, por estar traballando con datos trimestrais (se fosen anuais sería frecuencia igual a 12, que é a que mostra por defecto). Unha vez seleccionada a frecuencia que se axusta aos nosos datos temos un calendario dividido en trimestres, no que xa

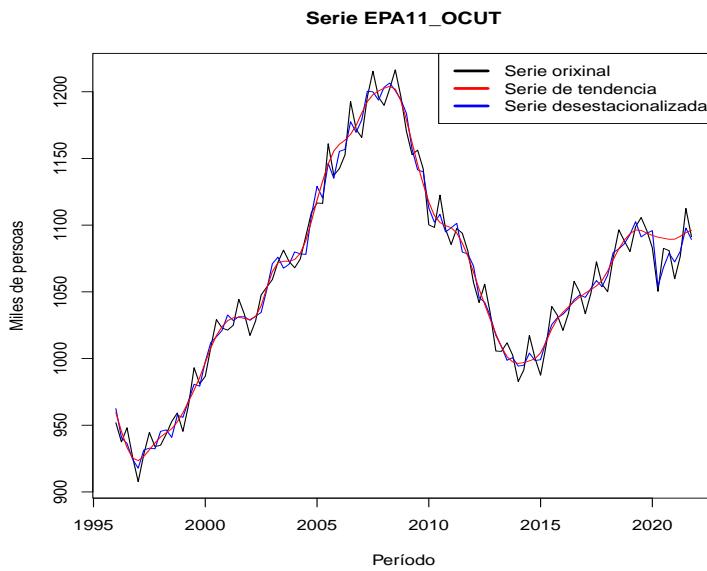


Figura 3.2: Serie orixinal de EPA11_OCUT, xunto coa serie correxida de estacionalidade e calendario, en cor azul, ademais da serie de tendencia en vermello. Axustes feitos para o modelo calculado por defecto por JDemetra+.

podemos establecer o atípico que queremos: un LS no segundo trimestre de 2020. Ademais desmarcamos a opción de selección de outliers automaticamente, para que o programa non busque máis atípicos, só queremos o atípico que acabamos de fixar (por simplicidade do modelo).

Análise gráfica e diagnose do novo modelo

Se agora executamos de novo o software de axuste obtemos un novo modelo, esta vez cun comportamento moito mellor graficamente no 2020, ante a aparición do COVID-19, como podemos apreciar na Figura 3.3.

En particular, con este novo modelo conseguimos corregir o aviso de problema de descomposición do modelo. Ademais, o axuste segue sendo moi bo a nivel de diagnose do modelo, xa que todos os tests devolven resultados favorables ás hipóteses de interese. En particular cúmprense as hipóteses de normalidade, independencia e linearidade dos residuos, como se mostra nos test do Cadro 3.1. Nótese que os números que aparecen entre parénteses nos distintos test de independencia e linealidade denotan o número de retardos que se teñen en conta para o cálculo do estadístico (véxase Sección 2.4).

Podemos mencionar o feito de que en JDemetra+ obtense unha pequena deficiencia nos test de estacionalidade dos residuos do último período, poñendo en dúbida a ausencia de estacionalidade para eles, pero como obtemos tanto test favorables como desfavorables, non está moi claro que sexa un factor realmente alarmante. O novo modelo segue un ARIMA(3,1,0)(0,1,1) sen media, igual que no primeiro intento, cun AICc de 758.79 empeorando o AICc=736.34 do modelo por defecto. Ademais temos certa información como pode ser que a serie foi transformada logaritmicamente e que non hai efectos nin de Working days (Trading days) nin de Semana Santa. En resumo, este novo modelo é sinxelo e axústase moi ben aos datos brutos, ademais correxe o problema de descomposición do modelo, entón, a pesar de que a nivel de AICc o modelo empeora lixeiramente, xa que, ten un AICc maior que o do modelo automático, temos argumentos para considerar que este segundo modelo é máis adecuado. Polo tanto, este será o modelo que consideraremos para modelizar a serie EPA11_OCUT.

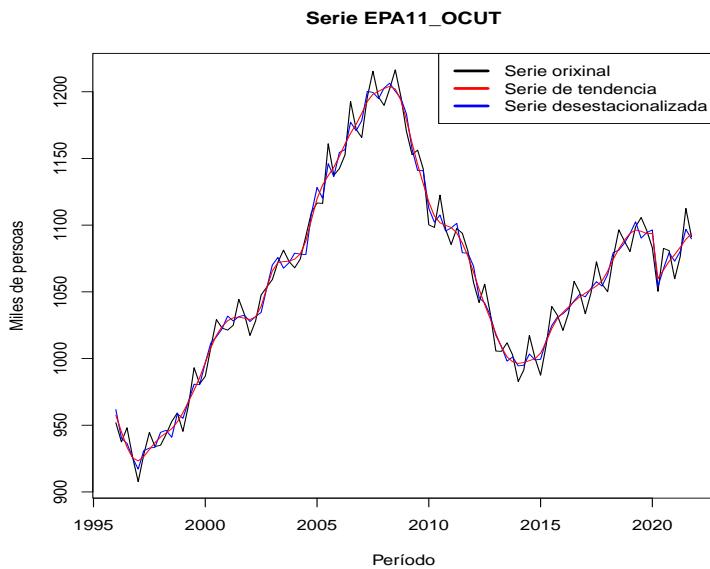


Figura 3.3: Serie orixinal de EPA11_OCUT, xunto coa serie correxida de estacionalidade e calendario, en cor azul, ademais da serie de tendencia en vermello. Axustes feitos para o modelo calculado por defecto por JDemetra+ incluíndo as modificacións do usuario.

3.2. Índice de producción industrial (IPI)

A continuación, cambiando un pouco a temática, pasaremos a estudar a modelización da serie de Índice xeral de producción industrial de Galicia, na base de 2015 (IPI15XE), que ten por fonte de información o Índice de producción industrial do INE.

O índice de producción industrial (IPI) é un indicador conxuntural cuxa finalidade é medir a evolución mensual da actividade produtiva das ramas industriais; é dicir, das industrias extractivas, manufactureiras e a subministración de enerxía eléctrica, gas, vapor e aire acondicionado. Mide, polo tanto, a evolución conxunta da cantidade e da calidade producidas, eliminando a influencia dos prezos.

En particular, trátase dunha serie de tempo mensual na que temos os datos agrupados por meses, de modo que hai 12 observacións por ano, con datos dispoñibles dende xaneiro do 2002 ata abril do 2022. Representamos a serie dos datos brutos na Figura 3.4.

Axuste automático

Ao executar o axuste da serie de tempo cos parámetros pre establecidos, de novo, na pestaña de resultados principais o primeiro que obtemos é a advertencia: “decomposition. Modelo decomposition: Parameters cut off”. Polo tanto, debemos modificar o axuste que se realizou sobre a serie de forma automática para corrixir o erro de descomposición.

Análise gráfica e diagnose do modelo

Podemos apreciar que os resultados de validación do modelo son favorables en xeral (a normalidade e independencia acéptanse pero están preto do límite de aceptación, como se pode ver no Cadro 3.2). Polo tanto, seguindo a idea da serie anterior, debemos modificar o modelo calculado automaticamente sen incluir cambios demasiado grandes para conseguir desfacermos do problema de descomposición sen perder os bons resultados de validación do modelo. Trataremos de facer pequenas modificacións sobre as especificacións devoltas polo programa.

<i>Normality of the residuals</i>	
	P-value
Mean	0.3841
Skewness	0.6932
Kurtosis	0.9092
Normality	0.8584

<i>Independence of the residuals</i>	
	P-value
Ljung-Box(16)	0.8518
Box-Pierce(16)	0.8982
Ljung-Box on seasonality(2)	0.8776
Box-Pierce on seasonality(2)	0.8849

<i>Linearity of the residuals</i>	
	P-value
Ljung-Box on squared residuals(16)	0.9234
Box-Pierce on squared residuals(16)	0.9419

Cadro 3.1: Tests para o modelo calculado por defecto por JDemetra+ coas modificacións incluídas polo usuario para a serie EPA11_OCUT.

O modelo obtido é un ARIMA(2,2,1)(0,1,1) sen media, e os atípicos consideran un AO no mes de xullo de 2008 e un TC no mes de abril de 2020. Ademais, visualmente na Figura 3.5 o axuste da serie é aceptable pero parece que a serie de tendencia podería ter un mellor comportamento a comezos do 2020, no período de aparición do COVID-19.

Novo modelo construído a partir da intervención do usuario

Comezaremos identificando os atípicos, ver Sección 2.1.2. Logo, como un TC no mes de abril de 2020 supón problemas na descomposición do modelo, podemos tratar de axustar a serie considerando no seu lugar un AO. Se levamos a cabo dita idea, continuando coa opción de selección de atípicos automática, agora resulta como modelo seleccionado o de liñas aéreas, ARIMA(0,1,1)(0,1,1), sen media, e os atípicos seleccionados son:

1. LS 11/2006
2. AO 7/2008

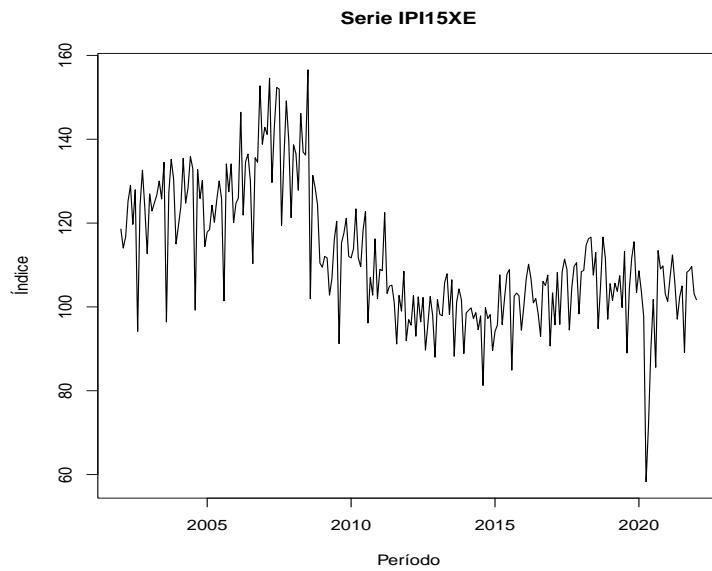


Figura 3.4: Serie orixinal de IPI15XE.

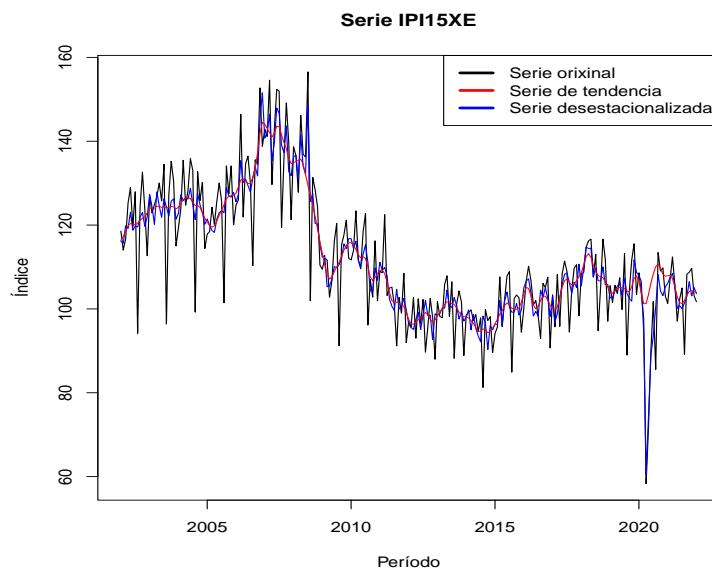


Figura 3.5: Serie orixinal de IPI15XE, xunto coa serie corregida de estacionalidade e calendario, en cor azul, ademais da serie de tendencia en vermello. Axustes feitos para o modelo calculado por defecto por JDemetra+.

3. TC 3/2020
4. AO 4/2020 (definido polo usuario)
5. AO 5/2020

Antes de analizar sequera a calidad do axuste debemos rexeitar sempre esta combinación de atípicos

<i>Normality of the residuals</i>	
	P-value
Mean	0.3605
Skewness	0.5945
Kurtosis	0.0111
Normality	0.0214

<i>Independence of the residuals</i>	
	P-value
Ljung-Box(24)	0.0138
Box-Pierce(24)	0.0255
Ljung-Box on seasonality(2)	1.0000
Box-Pierce on seasonality(2)	1.0000

<i>Linearity of the residuals</i>	
	P-value
Ljung-Box on squared residuals(24)	0.7243
Box-Pierce on squared residuals(24)	0.7763

Cadro 3.2: Tests para o modelo calculado por defecto por JDemetra+ para a serie IPI15XE.

producida no 2020 porque carece de lóxica. Non se poden considerar TC ou LS precedidos ou seguidos de varios AO, xa que, por definición de cada un destes procesos, ou se produce un ou o outro pero non todos xuntos. Considerar un TC seguido de dous AO podería indicar que algún atípico non logra captar completamente o efecto do suceso e polo tanto o programa intenta cubrilo engadindo outro atípico. En particular, por simplicidade, imos considerar o caso no que manualmente introducimos todos estes atípicos menos o TC de marzo de 2020, xa que un AO no mes de abril e outro no de maio podería ser razonable tendo en conta que no mes de abril a industria viuse obrigada a pechar ou reducir a súa producción ata niveis mínimos e no mes de maio rematou o confinamento e con el a producción repuntou. Ademais, graficamente tamén se aprecian os efectos do que podería ser un LS en novembro do 2006 e un AO en xullo de 2008. Polo tanto, agora sen considerar a selección automática de atípicos e introducindo manualmente os que acabamos de comentar, obtemos un ARIMA(2,1,0)(0,1,1) sen media cuns bos resultados en canto á diagnose do modelo. Con todo, como no axuste previo nos saiu por defecto o modelo de liñas aéreas podésemos ocorrer probar que resultados se obteñen se especificamos directamente que queremos dito modelo.

Normality of the residuals

	P-value
Mean	0.6471
Skewness	0.8091
Kurtosis	0.0487
Normality	0.0737

Independence of the residuals

	P-value
Ljung-Box(24)	0.1593
Box-Pierce(24)	0.2163
Ljung-Box on seasonality(2)	0.9252
Box-Pierce on seasonality(2)	1.0000

Linearity of the residuals

	P-value
Ljung-Box on squared residuals(24)	0.7259
Box-Pierce on squared residuals(24)	0.7820

Cadro 3.3: Tests para o modelo calculado por defecto por JDemetra+ xunto coas modificacións feitas polo usuario para a serie IPI15XE.

Análise gráfica e diagnose do novo modelo

Ao executar estas especificacións, é dicir, modelo ARIMA(0,1,1)(0,1,1) sen media (se o consideramos con media os resultados son en esencia os mesmos), con atípicos: LS en novembro de 2006, AO en xullo de 2008, AO en abril de 2020 e AO en maio de 2020, chegamos a un axuste considerablemente bo, onde graficamente mellorouse a serie da tendencia no período de aparición do COVID-19 (como se pode apreciar na Figura 3.6) e os test de validación son todos favorables, remarcando que os de normalidade a pesar de que tamén son favorables, están máis preto do límite de aceptación, como podemos ver nos test do Cadro 3.3. En canto ao criterio AICc, respalda a nosa elección, xa que no modelo inicial tiña un valor de 1386.42 fronte ao actual de 1370.91. Como información a maiores tense que non se considerou a transformación logarítmica da serie mentres que si se incluíron os regresores de Working days e Semana Santa no modelo final.

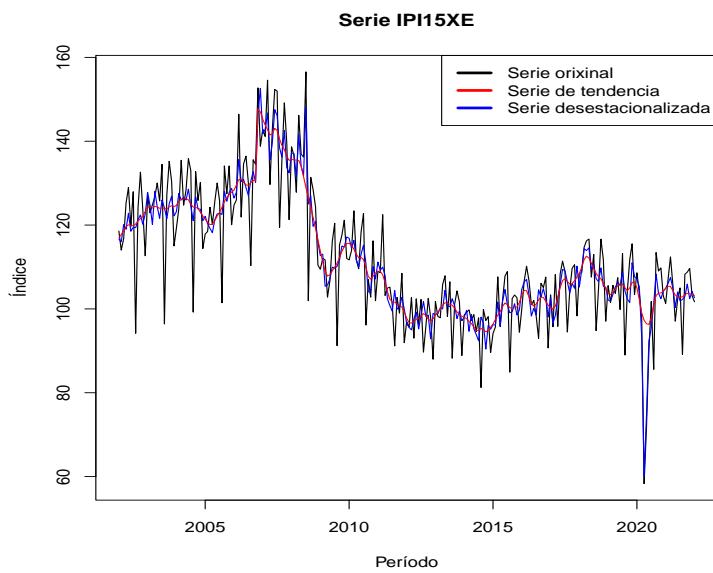


Figura 3.6: Serie orixinal de IPI15XE, xunto coa serie corregida de estacionalidade e calendario, en cor azul, ademais da serie de tendencia en vermello. Axustes feitos para o modelo calculado por defecto por JDemetra+ incluíndo as modificacíons do usuario.

3.3. Transporte marítimo.

A seguinte serie que imos analizar é a de transporte marítimo que inclúe o total de mercancías cargadas, descargadas e transbordadas en Galicia (TMM), cuxa información foi directamente subministrada polos Portos do Estado. Trátase dunha serie mensual que contén datos dende xaneiro do 1980 ata abril do 2022.

En particular, o tráfico portuario de mercadorías inclúe:

- Toneladas cargadas e descargadas a granel líquidos: as transportadas por un buque no mes de referencia. Os graneis líquidos son mercancías que non van envasadas ou empaquetadas e cuxo cargamento completo se considere unha unidade. Cárganse e descárganse por tubos.
- Toneladas cargadas e descargadas a granel sólidos: as transportadas por un buque no mes de referencia. Os graneis sólidos son mercancías que non van envasadas ou empaquetadas e cuxo cargamento completo se considere unha unidade. Cárganse e descárganse con grúas de culler ou por medios específicos, xeralmente de grande capacidade.
- Toneladas cargadas e descargadas de mercancía xeral: as transportadas por un buque no mes de referencia. A mercancía xeral son mercancías que van envasadas ou empaquetadas. Cárganse e descárganse con grúas.
- Toneladas descargadas de pesca fresca: as descargadas por buques pesqueiras, directamente na lonxa pesqueira para a súa poxa, no mes de referencia. Toneladas cargadas de avituallamento de produtos petrolíferos: as cargadas polos buques no mes de referencia, como combustible.
- Toneladas cargadas doutros avituallamentos: as cargadas polos buques no mes de referencia, necesarias para a navegación do buque e a súa vida nel, tales como víveres, auga potable, xeo, lubricantes, etc.

Na Figura 3.7 aparece representada a serie dos datos brutos.

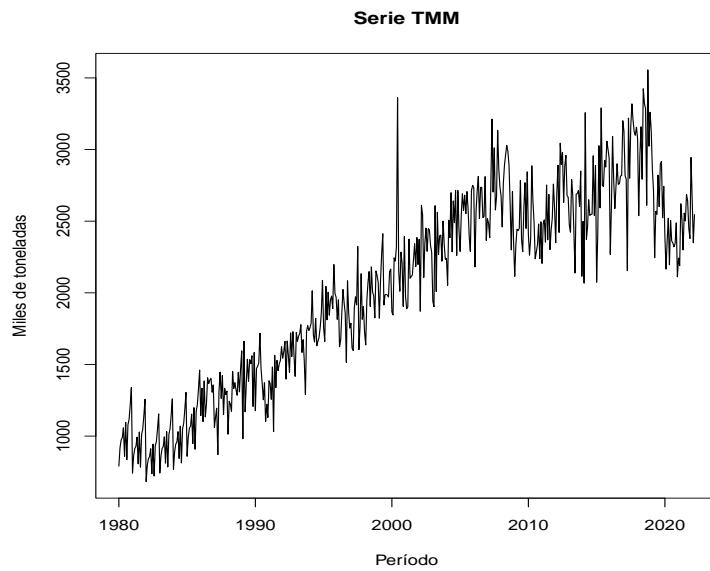


Figura 3.7: Serie orixinal de TMM.

Axuste automático

O primeiro axuste, realizado cos parámetros por defecto, estima un modelo ARIMA(2,1,1)(1,0,1) sen media cuxo valor de AICc é 6771.59. Indicásenos que os datos brutos foron transformados logarítmicamente e que non se engadiron ao modelo regresores de Working days (Trading days) nin de Semana Santa. Ademais, os atípicos seleccionados automaticamente foron: un TC en xaneiro de 1981, un TC en xaneiro de 1982 e un AO en xuño de 2000.

Análise gráfica e diagnose do modelo

Este modelo non será válido, xa que falla a hipótese de independencia, como podemos apreciar no Cadro 3.4, onde mostramos os p-valores de todos os test de independencia que se levan a cabo en JDmetra+ (cuxa hipótese nula é a independencia dos residuos). Os p-valores vermellos indican que os test son desfavorables, os verdes que son favorables e os amarelos que son favorables pero están preto do límite ou valor no que se rexeitaría a hipótese nula.

Por outra parte, visualmente, na Figura 3.8, podemos apreciar un axuste razoable.

Novo modelo construído a partir da intervención do usuario

A serie orixinal é moi errática, no sentido de que ten moitos atípicos (non é unha serie suave senón que ten moitos picos e saltos), polo que poderíamos plantexarnos dúas posibilidades, consideralos todos ou non considerar ningún.

Se comezamos polo caso máis fácil, que sería non considerar ningún atípico, seguimos sen ter independencia dos residuos, polo que ese caso non nos vale. Logo, comezamos un proceso de proba e análise no que imos modificando as posibles combinacións de atípicos e avaliando cada unha delas mediante os resultados obtidos graficamente e a diagnose do modelo. Despois deste proceso chegamos a unha combinación aceptable, considerando o modelo ARIMA(2,1,1)(1,0,1) sen media, xunto cos seguintes atípicos (Sección 2.1.2 para máis información sobre atípicos):

1. TC 1/1981
2. TC 1/1982

Lag	Autocorrelation	Standard deviation	Ljung-Box test	P-Value	Box-Piece test	P-Value
1	0.0048	0.0445				
2	0.0214	0.0445				
3	0.0228	0.0445				
4	-0.0585	0.0445				
5	0.0415	0.0445				
6	-0.0185	0.0445	3.3206	0.0684	3.2805	0.0701
7	0.0486	0.0445	4.5363	0.1035	4.4746	0.1067
8	-0.0307	0.0445	5.0242	0.1700	4.9529	0.1753
9	0.0543	0.0445	6.5519	0.1616	6.4475	0.1681
10	-0.0227	0.0445	6.8192	0.2344	6.7085	0.2432
11	0.1251	0.0445	14.9411	0.0207	14.6226	0.0234
12	-0.0429	0.0445	15.8991	0.0261	15.5542	0.0295
13	-0.0239	0.0445	16.1975	0.0396	15.8438	0.0447
14	0.0119	0.0445	16.2716	0.0614	15.9155	0.0687
15	-0.1927	0.0445	35.7123	0.0001	34.7057	0.0001
16	-0.0310	0.0445	36.2180	0.0002	35.1934	0.0002
17	0.0135	0.0445	36.3139	0.0003	35.2858	0.0004
18	-0.0223	0.0445	36.5759	0.0005	35.5374	0.0007
19	-0.0138	0.0445	36.6764	0.0008	35.6338	0.0012
20	-0.0129	0.0445	36.7643	0.0014	35.7179	0.0019
21	-0.0340	0.0445	37.3776	0.0019	36.3034	0.0026
22	0.0184	0.0445	37.5574	0.0028	36.4747	0.0040
23	-0.1057	0.0445	43.5047	0.0007	42.1293	0.0011
24	0.0235	0.0445	43.7985	0.0010	42.4081	0.0016

Cadro 3.4: Test completos de independencia da serie TMM sobre o axustes feitos para o modelo calculado por defecto por JDemetra+.

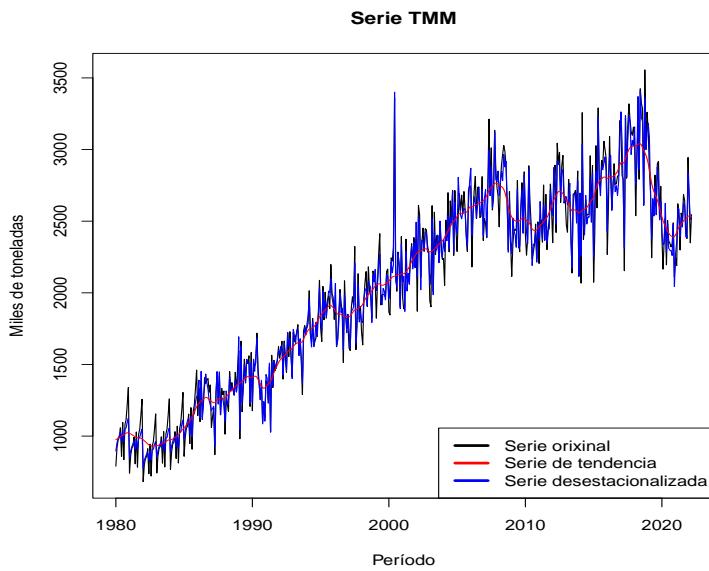


Figura 3.8: Serie orixinal de TMM, xunto coa serie correxida de estacionalidade e calendario, en cor azul, ademais da serie de tendencia en vermello. Axustes feitos para o modelo calculado por defecto por JDemetra+.

3. TC 1/1984
4. AO 4/1987
5. AO 2/1989
6. AO 6/2000
7. AO 3/2014
8. AO 10/2018
9. AO 4/2019

e sen a opción de selección automática de atípicos. Esta combinación acada uns resultados razonables, cun $AICc=6735.57$, mellorando o $AICc$ do axuste automático, onde todos os tests son favorables. En particular, a normalidade, a independencia e a linearidade están preto do límite de aceptación de ditas hipóteses. Visualmente tampouco temos ningún argumento en contra deste modelo. Con todo, acabamos de engadir de xeito manual nove atípicos, o que fai que o modelo non sexa tan manexable como podería ser outro con un par de atípicos. Tendo en conta este desexo de simplificar o modelo, sacrificaremos algunha das características obtidas con este modelo para obter un modelo máis simple e manexable. Isto lévanos a considerar soamente dous atípicos: un AO en abril de 1987 e un AO en febreiro de 1989. Se agora lles aplicamos a selección de modelo automático, obtemos un ARIMA(0,1,2)(0,1,1) sen media, cuxa serie orixinal foi transformada logarítmicamente e non inclúe regresores de Working days (Trading days) nin de Semana Santa.

Análise gráfica e diagnose do novo modelo

Con este último modelo obtemos un $AICc$ de 6652.29, menor que os $AICc$ obtidos cos dous modelos anteriores, polo tanto, este criterio é un motivo engadido para decantarnos polo modelo máis simple. Pero antes de tomar a decisión final debemos analizar os resultados do axuste gráfico, que como se pode apreciar na Figura 3.9, devolve series que teñen un comportamento aceptable con respecto á serie orixinal. En canto aos test de diagnose do modelo, como podemos ver no Cadro 3.5, que contén a versión extendida

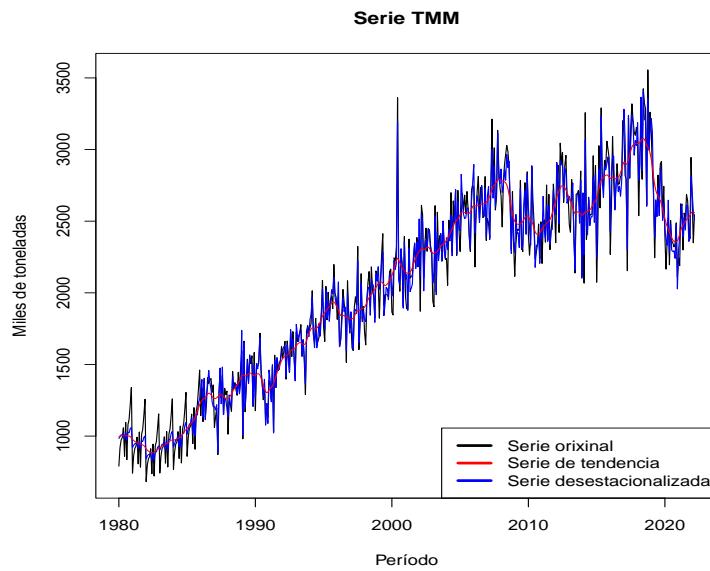


Figura 3.9: Serie orixinal de TMM, xunto coa serie correxida de estacionalidade e calendario, en cor azul, ademais da serie de tendencia en vermello. Axustes feitos para o modelo calculado por defecto por JDemetra+ incluíndo as modificacións do usuario.

dos test de independencia, corríxese parcialmente o problema de non independencia dos residuos, aínda que sigue habendo moitos test amarelos (no límite de aceptación).

Ademais, como se pon de manifesto nos test do Cadro 3.6, non temos normalidade nin linearidade dos residuos (nótese que se consideramos o modelo con nove atípicos en vez de dous si teríamos estas hipóteses, con tests en amarelo).

Polo tanto, se buscamos un modelo simple que axuste ben os datos e que sexa razonable en canto a diagnose, como é o noso caso, podemos aceptar este último modelo. Se noutro caso precisasemos do cumprimento das hipóteses de normalidade e linearidade simplemente teríamos que pasar a considerar o modelo con nove atípicos. Para o problema que nos atinxé na modelización de series de tempo de convivencia de Galicia, decantáremos polo modelo máis simple, xa que atendendo a todos os criterios de diagnose do modelo, como son os test de validación, o axuste gráfico e o valor do AICc, as diferenzas non son suficientemente significativas como para engadir complexidade ao modelo.

3.4. Ocupación hoteleira: viaxeiros entrados.

Agora, pasando ao sector do turismo, veremos o que pasa se tratamos de modelizar a serie de ocupación nos establecementos hoteleiros contabilizando o número de viaxeiros entrados (EOHVE_04), cuxa información ten por fonte a Enquisa continua de ocupación hoteleira do INE. Comprende a todas aquelas persoas que realizan unha ou máis noites seguidas no mesmo aloxamento. Trátase dunha serie mensual con datos dende xaneiro do 1999 ata abril do 2022. Representamos a serie dos datos brutos na Figura 3.10.

Axuste automático

Segundo o procedemento estándar que levamos a cabo coas outras series, comezamos executando o axuste da serie cos parámetros por defecto. O resultado é un modelo de liñas aéreas con dez atípicos.

Os atípicos que foron engadidos ao modelo de forma automática son:

Lag	Autocorrelation	Standard deviation	Ljung-Box test	P-Value	Box-Piece test	P-Value
1	-0.0080	0.0450				
2	0.0784	0.0450				
3	-0.0345	0.0450				
4	-0.0650	0.0450	5.7951	0.0161	5.7387	0.0166
5	-0.0232	0.0450	6.0652	0.0482	6.0050	0.0497
6	-0.0218	0.0450	6.3028	0.0978	6.2388	0.1006
7	0.0665	0.0450	8.5257	0.0741	8.4213	0.0773
8	-0.0578	0.0450	10.2093	0.0695	10.0710	0.0732
9	0.0304	0.0450	10.6763	0.0989	10.5277	0.1041
10	-0.0315	0.0450	11.1800	0.1310	11.0191	0.1378
11	0.0836	0.0450	14.7244	0.0647	14.4706	0.0703
12	-0.0251	0.0450	15.0438	0.0897	14.7810	0.0971
13	-0.0267	0.0450	15.4068	0.1179	15.1331	0.1273
14	0.0424	0.0450	16.3251	0.1295	16.0217	0.1403
15	-0.1541	0.0450	28.4695	0.0047	27.7499	0.0060
16	-0.0136	0.0450	28.5645	0.0075	27.8414	0.0095
17	0.0178	0.0450	28.7269	0.0114	27.9977	0.0142
18	-0.0023	0.0450	28.7296	0.0174	28.0002	0.0216
19	0.0147	0.0450	28.8409	0.0250	28.1068	0.0307
20	-0.0557	0.0450	30.4434	0.0233	29.6382	0.0291
21	-0.0075	0.0450	30.4727	0.0331	29.6662	0.0408
22	-0.0099	0.0450	30.5240	0.0455	29.7150	0.0555
23	-0.0739	0.0450	33.3676	0.0307	32.4152	0.0391
24	0.0379	0.0450	34.1160	0.0352	33.1244	0.0449

Cadro 3.5: Test de independencia da serie TMM sobre o axustes feitos para o modelo calculado por defecto por JDemetra+ incluíndo as modificacións do usuario.

Normality of the residuals

	P-value
Mean	0.8391
Skewness	0.6259
Kurtosis	0.0000
Normality	0.0000

Independence of the residuals

	P-value
Ljung-Box(24)	0.0352
Box-Pierce(24)	0.0449
Ljung-Box on seasonality(2)	0.6878
Box-Pierce on seasonality(2)	1.0000

Linearity of the residuals

	P-value
Ljung-Box on squared residuals(24)	0.0013
Box-Pierce on squared residuals(24)	0.0019

Cadro 3.6: Resultados dos test encargados da análise dos residuos da serie TMM sobre o axustes feitos para o modelo calculado por defecto por JDemetra+ incluíndo as modificacións do usuario.

1. TC 9/1999
2. LS 11/2004
3. AO 4/2017
4. TC 7/2018
5. LS 7/2019
6. LS 3/2020
7. LS 4/2020
8. LS 11/2020
9. AO 1/2021
10. TC 4/2021

Análise gráfica e diagnose do modelo

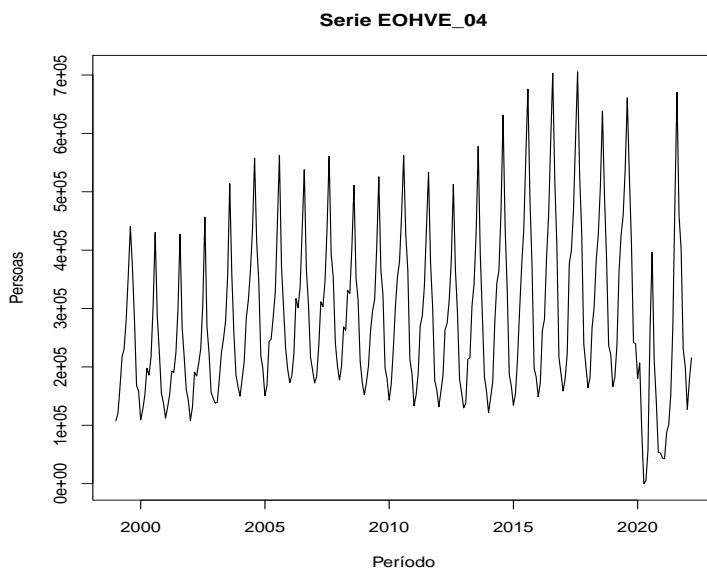


Figura 3.10: Serie orixinal de EOHVE_04.

Como se aprecia na Figura 3.11, este modelo graficamente presenta un comportamento “estraño” ou esaxerado no período de aparición do COVID-19. O valor do AICc para este modelo é de 6004.76. En canto aos diferentes test de diagnose, obtemos resultados aceptablemente bos (con algún test en amarelo), polo que de novo, debemos conseguir un mellor axuste gráfico da serie facendo pequenas modificacíons sobre o modelo dado automaticamente para conservar os bos resultados de validación.

Novo modelo construído a partir da intervención do usuario

Na Figura 3.11 apreciabamos que a serie ten un comportamento “similar” ao longo dos anos pero ao chegar ao ano 2020 a aparición do COVID-19 provoca un cambio neste. Polo tanto, por simplicidade descartaremos os atípicos todos que se consideraron de forma automática para os anos que preceden ao 2020. En canto aos dous cambios de nivel considerados en marzo e abril do 2020, preferiremos unha combinación do tipo: un AO en marzo do 2020 e un LS en abril do 2020. No ano 2021, os atípicos considerados non acaban de representar o comportamento real da serie. Se nos paramos a analizar un pouco o que ocorre nese ano, en xaneiro do 2021 si que se produce un cambio de nivel, pero aproximadamente en xuño ten lugar outro cambio de nivel, este segundo levando a serie de novo aos niveis prepandemia. Entón, se consideramos esta combinación de atípicos: AO en marzo de 2020, LS en abril de 2020, LS en novembro de 2020, LS en xaneiro de 2021 e LS en xuño de 2021, obtemos un modelo automático ARIMA(1,0,0)(1,1,1) sen media, con axuste gráfico e diagnose aceptables, salvo polo feito de que non temos normalidade nos residuos, como se pode apreciar nos test do Cadro 3.7. O valor do AICc é de 6086.14.

Con todo, seguimos tendo un número relativamente elevado de atípicos, polo que quizais podemos atopar un modelo máis simple que devolva un axuste razonable. Pois ben, o gran impacto do COVID-19 prodúcese en 2020, en particular no mes de marzo foi cando se estableceu o confinamento e no mes de maio cando se deu por finalizado, polo tanto, sería bastante lóxico tratar de modelar este fenómeno con dous cambios de nivel, un en marzo e outro en maio. Isto será o que faremos a continuación, considerar únicamente como atípicos un LS no mes de marzo de 2020 e outro no mes de maio.

Con estas especificacións obtemos un ARIMA(0,1,0)(0,1,1) sen media con regresor de Semana Santa e sen regresor de Working days (Trading days), cuxa serie orixinal non se lle aplicou a transformación logarítmica.

Normality of the residuals

	P-value
Mean	0.0341
Skewness	0.0291
Kurtosis	0.0000
Normality	0.0000

Independence of the residuals

	P-value
Ljung-Box(24)	0.1679
Box-Pierce(24)	0.2188
Ljung-Box on seasonality(2)	0.8294
Box-Pierce on seasonality(2)	0.8375

Linearity of the residuals

	P-value
Ljung-Box on squared residuals(24)	0.1525
Box-Pierce on squared residuals(24)	0.1902

Cadro 3.7: Resultados dos test encargados da análise dos residuos da serie EOHVE_04 sobre o axuste feito para o modelo calculado por defecto por JDemetra+ incluíndo modificacións de proba ou intermedias dadas polo usuario.

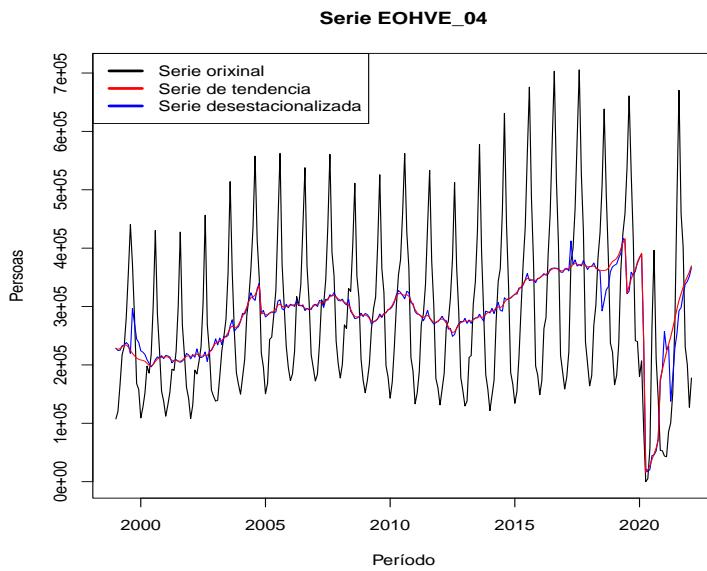


Figura 3.11: Serie orixinal de EOHVE_04, xunto coa serie correxida de estacionalidade e calendario, en cor azul, ademais da serie de tendencia en vermello. Axustes feitos para o modelo calculado por defecto por JDemetra+.

Análise gráfica e diagnose do novo modelo

O valor do criterio AICc para o novo modelo é de 6181.44, polo que neste caso non respalda a nosa idea de buscar o modelo máis sinxelo, xa que obtivemos valores de AICc menores para os dous modelos anteriores. Con todo, como as diferenzas non son moi significativas, seguiremos considerando este modelo como adecuado. En particular, na Figura 3.12 podemos apreciar un mellor comportamento das series correxidas no período do COVID-19. En canto á diagnose do modelo, de novo non temos normalidade dos residuos, polo resto obténense resultados aceptablemente bos (como podemos ver no Cadro 3.8). Polo tanto, escolleremos como modelo de axuste para a serie de tempo EOHVE_04 este último polo equilibrio que establece entre resultados favorables e simplicidade do modelo. É dicir, tanto a nivel de test de diagnose, axuste gráfico e valor do AICc, o modelo máis complexo (a nivel de número de atípicos) non supón melloras suficientemente significativas, polo que o descartamos.

3.5. Xeración de enerxía eléctrica.

Por último consideraremos a serie de tempo de producción de enerxía eléctrica (GEETO), que se obtivo a partir do balance eléctrico da Rede eléctrica de España.

A xeración de enerxía eléctrica recolle toda as transformación de enerxía do tipo: química, cinética, térmica, lumínica, nuclear, solar, etc, en enerxía eléctrica. Trátase dunha serie de datos mensuais, que comprende datos dende xaneiro do 2013 ata febreiro do 2022. Debemos mencionar que as series de producción enerxética teñen unha naturaleza moi errática, polo que o axuste gráfico non vai ser tan certeiro como nos gustaría. A representación gráfica dos datos brutos desta serie de tempo pódese observar na Figura 3.13.

Axuste automático

Procedendo como coas outras series, comezamos polo axuste da serie coas preespecificacións por defecto. Pero neste caso JDemetra+ fai un bo traballo e devólvenos un modelo cun comportamento bastante razonable, polo que non seguiremos buscando.

Normality of the residuals

	P-value
Mean	0.7087
Skewness	0.0000
Kurtosis	0.0000
Normality	0.0000

Independence of the residuals

	P-value
Ljung-Box(24)	0.1246
Box-Pierce(24)	0.1713
Ljung-Box on seasonality(2)	0.2843
Box-Pierce on seasonality(2)	0.3036

Linearity of the residuals

	P-value
Ljung-Box on squared residuals(24)	0.1671
Box-Pierce on squared residuals(24)	0.1985

Cadro 3.8: Resultados dos test encargados da análise dos residuos da serie EOHVE_04 sobre o axuste feito para o modelo calculado por defecto por JDemetra+ incluíndo as modificacións finais do usuario.

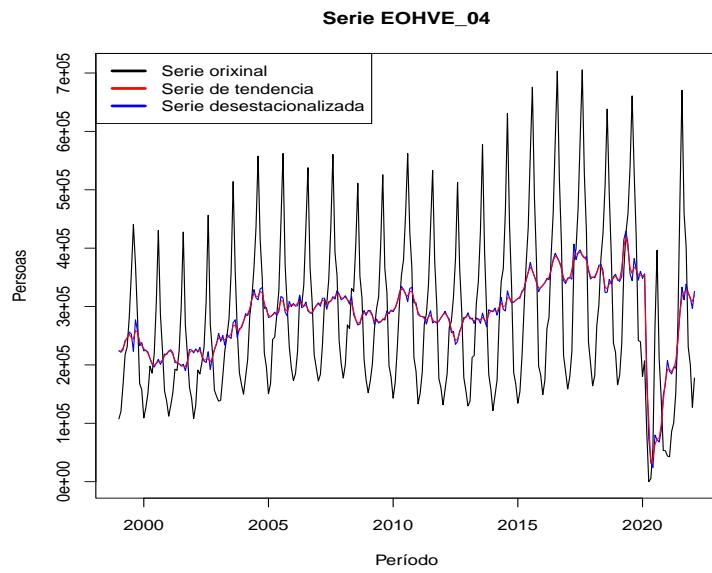


Figura 3.12: Serie orixinal de EOHVE.04, xunto coa serie correxida de estacionalidade e calendario, en cor azul, ademais da serie de tendencia en vermello. Axustes feitos para o modelo calculado por defecto por JDemetra+ incluíndo as modificacións do usuario.

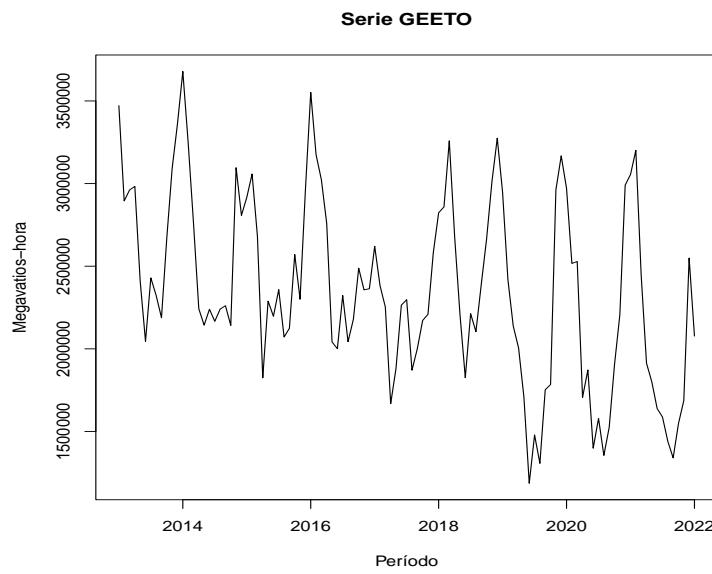


Figura 3.13: Serie orixinal de GEETO.

O axuste en cuestión considera un modelo ARIMA(1,0,0)(0,1,1) con media, sen atípicos, cuxa serie orixinal foi transformada con logaritmos e non inclúe regresores nin de Working days (Trading days) nin de Semana Santa.

Análise gráfica e diagnose do modelo

A nivel gráfico, xa adiantabamos que a serie ten moitas perturbacións polo que obter series correxidas que se aproximen “moito” á serie orixinal é unha tarefa complicada. Sen embargo o axuste obtido capta de forma razonable a tendencia e as subidas e baixadas da serie orixinal, como se pode apreciar na Figura 3.14. Ademais os tests de validación e diagnose do modelo devolven resultados favorables en xeral. Os resultados principais que imprime JDmetra+ pódense ver no Cadro 3.9, onde todos os test aparecen en cor verde (favorables).

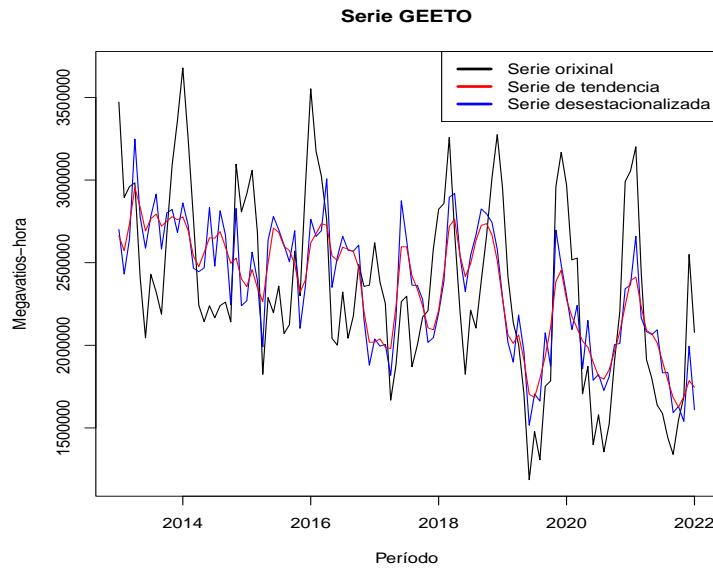


Figura 3.14: Serie orixinal de GEETO, xunto coa serie correxida de estacionalidade e calendario, en cor azul, ademais da serie de tendencia en vermello. Axustes feitos para o modelo calculado por defecto por JDmetra+.

<u>Diagnostics</u>	
summary	Good
regarima residuals	
normality	Good (0.816)
independence	Good (0.663)
linearity	Good (0.358)
combined seasonality test	
Kruskall-Wallis test	Good (0.000)
test for the presence of seasonality assuming stability	Good (0.000)
evolutive seasonality test	Good (0.7419)
residual seasonality tests	
qs test on sa	Good (1.000)
qs test on i	Good (1.000)
f-test on sa (seasonal dummies)	Good (1.000)
f-test on i (seasonal dummies)	Good (1.000)
Residual seasonality (entire series)	Good (1.000)
Residual seasonality (last 3 years)	Good (1.000)
residual trading days test	
f-test on sa (td)	Good (0.080)
f-test on i (td)	Good (0.108)

Cadro 3.9: Resultados dos test de validación do modelo da serie GEETO sobre o axuste feito para o modelo calculado por defecto por JDemetra+.

Capítulo 4

Automatización do proceso en R

O programa JDemetra+ permítenos obter os modelos resultantes do proceso de desestacionalización para as series de tempo de conxuntura da economía galega seguindo un proceso bastante manual. Con todo, existen programas e linguaxes de programación más manexables que poderían ser de maior interese á hora de modelizar as series de tempo, en particular, centrarémonos na linguaxe de programación R. O noso obxectivo será levar á práctica as técnicas vistas para correxir as series temporais con JDemetra+, agora empregando R Project, de modo que poidamos obter un procedemento automático de axuste das series.

Para levar a cabo este procedemento precisamos instalar e cargar dúas librerías en R, que nos permitirán acceder a unha adaptación da maioría de funcións e resultados dos que dispón JDemetra+. Ditas librerías son: *RJDemetra* e *rjd3modelling* (que á súa vez conleva a necesidade de cargar a librería *rjd3toolkit* (Palate, 2022b)). *RJDemetra* (Quartier-la-Tente et al., 2022) é unha librería que nos achega unha interface arredor de JDemetra+, ofrecendo acceso á gran maioría de opcións e resultados de JDemetra+, incluíndo os dous métodos principais de axuste estacional xa mencionados, TRAMO-SEATS e X-13ARIMA-SEATS. Por outro lado, a librería *rjd3modelling* (Palate, 2022a) actúa como soporte de actualización resolvendo algunas das limitacións que presenta *RJDemetra*, por exemplo, *RJDemetra* non ten implementada a posibilidade de introducir un calendario de festividades particular, mentres que *rjd3modelling* contén unha serie de funcións que permiten construir calendarios específicos, que posteriormente poden ser engadidos ao procedemento de axuste estacional mediante a condición de variables definidas polo usuario ('usrdef.var').

RJDemetra inclúe, a maiores das ferramentas para proceder de xeito análogo a como se faría no propio software de JDemetra+, unha serie de funcións que permiten acceder aos elementos dun axuste estacional previamente realizado na interface de JDemetra+, as cales poden ser de utilidade se se deseja modelar as series en JDemetra+ e a continuación levar os resultados a R Project. Nótese que nos referiremos a un obxecto multi-processing como o resultado do axuste estacional dunha única serie, é dicir, en JDemetra+, coa opción de axuste multi-processing permítense modelizar varias series de forma conxunta, pero este caso non será do noso interese, polo que, aínda que falemos de obxectos multi-processing, neste traballo sempre nos estaremos a referir ao obxecto resultante do axuste estacional dunha única serie (considerando o calendario particular de Galicia). Ademais, os arquivos nos que se gardan os resultados do modelado das series de tempo chámense *workspace* e dispoñemos dunha serie de funcións para traballar con eles, que se presentan a continuación:

- *compute()*: carga un workspace multi-processing.
- *count()*: conta o número de obxectos dentro dun workspace ou multi-processing.
- *get_model()*: obtén o modelo axustado estacionalmente dun workspace.
- *get_name()*: obtén o nome Java dun multiprocessing.
- *get_object()*: obtén a serie de tempo cos datos brutos.
- *load_workspace()*: carga un workspace 'JDemetra+'.

- *new_workspace()*: crea un workspace ou un multi-processing.
- *save_workspace()*: garda un workspace.

Por exemplo, no noso caso, o seguinte código de R permítenos obter os resultados dos multi-processing levados a cabo para as series de tempo de interese. Comézase cargando o obxecto que contén os resultados, sobre o que se executa o cálculo do axuste estacional e por último pedímoslle que nos devolva os modelos finais para as series (gardado baixo o nome ‘modelo’), que será unha lista con 35 elementos, cada un dos cales corresponde ao axuste estacional dunha das 35 series que axustamos de forma individual.

```
exemplo<-load_workspace ("D:/TFM/R/WOrkspace/JDmetra_Alba.xml")
compute(exemplo)
modelo<-get_model(exemplo) # Modelo para cada Multi-processing
```

4.1. Calendario de Galicia

Unha vez cargadas as librerías, de forma análoga a como procedemos en JDemetra+, debemos construír o noso calendario particular de festividades de Galicia. O primeiro paso é crear un obxecto calendario coa función *calendar.new()*. A continuación a función *calendar.holiday()* permítenos definir certas festividades preestablecidas no programa por ser das más coñecidas. No noso caso, con este comando introduciremos: Aninovo (1 de xaneiro), Día internacional dos traballadores (1 de maio), Asunción (15 de agosto), Día de todos os santos (1 de novembro) e Nadal (25 de decembro). O resto de festivos non veñen incluídos na lista dos predifinidos polo programa, polo que empregaremos a función *calendar.fixedday()* para incluílos, entre os cales temos: Día de Reis (6 de xaneiro), Día das letras galegas (17 de maio), Día nacional de Galicia (25 de xullo), Festa nacional de España (12 de outubro), Día da Constitución (6 de decembro) e Día da Inmaculada Concepción (8 de decembro). Por último fáltanos incluír as festividades de San Xosé (19 de marzo) e San Xoán (24 de xuño), que non se poden introducir da mesma maneira, xa que, como xa comentamos previamente, ditos festivos non se producen todos os anos. En particular, existen dúas maneiras de proceder. A primeira sería coa función *calendar.fixedday()* indicándolle os valores que toman os parámetros ‘start=’ e ‘end=’ en formato data. Por exemplo, para crear un calendario que incluíse únicamente o festivo de San Xoán de 2022, escribiríamos:

```
GaliciaCalendar <- calendar.new()
calendar.fixedday(GaliciaCalendar, month = 6, day = 24, start="2022-01-01",
                   end="2022-12-31")
```

Con todo, no paquete *rjd3modelling*, existe unha función específica para aqueles festivos que só se producen unha vez, *calendar.singledate()*, polo que poderíamos introducir o festivo de San Xoán de 2022 de forma equivalente como segue:

```
calendar.singledate(GaliciaCalendar, "2022-06-24")
```

Unha vez feito isto, temos un calendario creado, pero dito calendario non ten un formato aplicable en *RJDmetra*, polo que teremos que traballar un pouco máis sobre el para poder incluílo á hora de realizar o axuste estacional. Nesta parte aparece unha das limitacións presentes aínda en *RJDmetra*, que será o feito de que só sexa capaz de detectar o efecto dun calendario particular mediante a especificación da variable definida polo usuario (‘usrdef.var’), polo que apareceran algunas pequenas discrepancias nos resultados, xa que en JDemetra+ o calendario non se introduce como ‘usrdef.var’.

Tal e como se indicou no capítulo de metodoloxía, construiremos os regresores de calendario incluíndolle o efecto de ano bisiesto (Leap Year). Ademais, debemos formular dúas modalidades de regresores a considerar, xa que, en JDemetra+, por defecto, realiza un test para decidir se se consideraba o efecto do calendario mediante

Working days, Trading days ou non se consideraba dito efecto, sen embargo, como nós temos que introducir o calendario como variable definida polo usuario, temos que darlle un formato previo, é dicir, decidir se o incluímos en formato Working days ou Trading days, onde por defecto tenderase a considerar coa estructura de Working days por ser máis sinxela. Con todo, é recomendable comparar os resultados cos obtidos considerando Trading days. Polo tanto, construiremos os regresores para ambos casos, Working days e Trading days. Cunha frecuencia de 12 meses e considerando un calendario que comeza en 1960 e remata en 2099, a continuación presentamos a estructura de ambos regresores para os doce primeiros meses considerados, no Cadro 4.1 o regresor de Working days e no Cadro 4.2 o de Trading days.

	week.Days	leap_year
1960 Jan	-6.00	0.00
1960 Feb	1.00	0.75
1960 Mar	20.50	0.00
1960 Apr	-1.50	0.00
1960 May	1.00	0.00
1960 Jun	12.00	0.00
1960 Jul	-5.00	0.00
1960 Aug	2.00	0.00
1960 Sep	2.00	0.00
1960 Oct	-5.00	0.00
1960 Nov	3.50	0.00
1960 Dec	0.00	0.00

Cadro 4.1: Regresor de Working days.

O código completo de como se constrúen ditos regresores pódese consultar no Apéndice B.

Cabe mencionar que non se realiza unha corrección por media a longo prazo no calendario, xa que, áinda que desde o INE se recomenda facela sempre para que a parte do efecto de calendario que contén o efecto media sexa asociado á compoñente de tendencia, no noso caso ao incluír as festividades de San Xosé e San Xoán que non se producen todos os anos debemos desestimar esta corrección para evitar que nos leve a resultados enganosos. É dicir, por defecto considerarase a corrección por media a longo prazo, pero cando no calendario aparece un evento que ocorre poucas veces, débense evitar correccións engañosas a longo plazo, como é o noso caso.

4.2. Especificacións JDemetra+

O seguinte paso consiste en introducir todas as especificacións que JDemetra+ emprega, por defecto, como punto de partida para calcular os modelos de axuste estacional, que se presentan nun panel con diferentes

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	leap_year
1960 Jan	-1.00	-1.00	-2.00	-1.00	-1.00	0.00	0.00
1960 Feb	1.00	0.00	0.00	0.00	0.00	0.00	0.75
1960 Mar	7.00	8.00	8.00	8.00	7.00	7.00	0.00
1960 Apr	0.00	0.00	0.00	0.00	1.00	1.00	0.00
1960 May	1.00	0.00	0.00	0.00	0.00	0.00	0.00
1960 Jun	4.00	4.00	5.00	5.00	4.00	4.00	0.00
1960 Jul	-2.00	-1.00	-1.00	-1.00	0.00	0.00	0.00
1960 Aug	0.00	1.00	1.00	0.00	0.00	0.00	0.00
1960 Sep	0.00	0.00	0.00	1.00	1.00	0.00	0.00
1960 Oct	0.00	-1.00	-2.00	-1.00	-1.00	0.00	0.00
1960 Nov	1.00	1.00	2.00	1.00	1.00	1.00	0.00
1960 Dec	1.00	0.00	1.00	1.00	2.00	2.00	0.00

Cadro 4.2: Regresor de Trading days.

apartados como o que se pode observar na Figura 4.1. En particular, moitos destes parámetros xa non serán modificados, polo que podemos construír unha función en R de modo que os elementos de entrada sexan os únicos que se modificarán. Polo tanto, intuitivamente os elementos de entrada serán os parámetros relacionados cos atípicos e co modelo ARIMA, así como os que permiten incluír os regresores do efecto de calendario. A función para introducir as especificacións en R é *tramoseats_spec()* (porque estamos realizando o axuste empregando o procedemento TRAMO-SEATS, mentres que con X-13ARIMA-SEATS sería *x13_spec()*). Para consultar o código co que se definen as diferentes especificacións máis polo miúdo véxase o Apéndice B.

Podemos dar unha breve definición do que significa cada un dos parámetros de entrada da función de especificacións que son: ‘outlier.enabled’, ‘usrdef.outliersEnabled’, ‘usrdef.outliersType’, ‘usrdef.outliersDate’, ‘usrdef.varEnabled’, ‘usrdef.var’, ‘easter.type’, ‘automdl.enabled’, ‘arima.mu’, ‘arima.p’, ‘arima.d’, ‘arima.q’, ‘arima.bp’, ‘arima.bd’ e ‘arima.bq’.

- **outlier.enabled:** parámetro de tipo lóxico que toma valores TRUE ou FALSE. Se é verdadeiro, permítense a detección automática de valores atípicos. Por defecto considérase ‘outlier.enabled = FALSE’.
- **usrdef.outliersEnabled:** parámetro de tipo lóxico que toma valores TRUE ou FALSE. Se é verdadeiro, o programa usa os valores atípicos especificados previamente polo usuario. Por defecto considérase ‘usrdef.outliersEnabled = TRUE’.
- **usrdef.outliersType:** vector que define o tipo de valores atípicos definidos polo usuario. Existen catro tipos de atípicos posibles: “AO” (aditivo), “LS” (cambio de nivel), “TC” (cambio transitorio) e “SO” (valor atípico estacional).

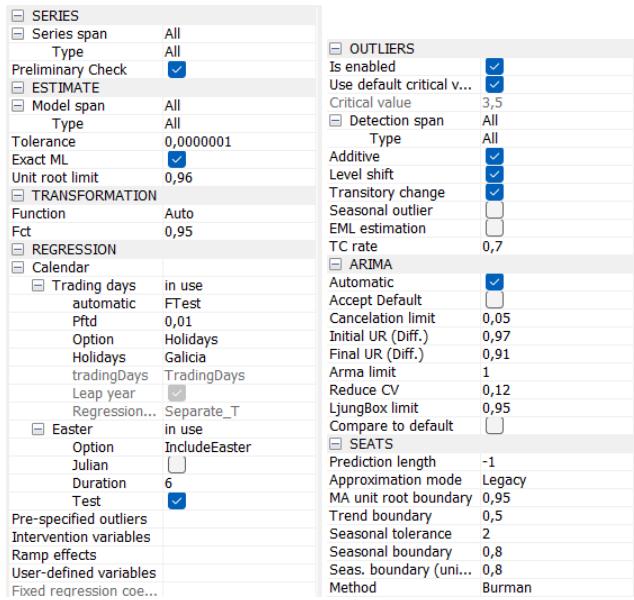


Figura 4.1: Panel de especificacións de JDemetra+.

- usrdef.outliersDate: vector que define as datas dos valores atípicos. As datas deben estar en formato data, “AAAA-MM-DD”.
- usrdef.varEnabled: parámetro de tipo lóxico que toma valores TRUE ou FALSE. Se é verdadeiro, o programa usa as variables definidas polo usuario. Por defecto considérase ‘usrdef.varEnabled = TRUE’.
- usrdef.var: serie temporal (ts) ou matriz de series temporais coas variables definidas polo usuario.
- easter.type: parámetro que especifica a presenza e a duración do efecto de Pascua. Podemos diferenciar entre: “Unused” se non se considera o efecto de Pascua; “Standard” cando inflúe no período de n días estritamente antes do Domingo de Pascua; “IncludeEaster” cando inflúe en todo o período (n) ata o Domingo de Pascua incluído; “IncludeEasterMonday” cando inflúe en todo o período (n) ata o luns de Pascua, incluído. Por defecto considérase ‘easter.type = “IncludeEaster” ’.
- automdl.enabled: parámetro de tipo lóxico que toma valores TRUE ou FALSE. Se é verdadeiro, o cálculo automático do modelo ARIMA está habilitado. En caso contrario, se é falso, pódense especificar os parámetros do modelo ARIMA. Por defecto considérase ‘automdl.enabled = FALSE’.
- arima.mu: parámetro de tipo lóxico que toma valores TRUE ou FALSE. Se é verdadeiro, a media considerase parte do modelo ARIMA. Por defecto considérase ‘arima.mu = FALSE’.
- arima.p: valor numérico. Denota a orde do polinomio autorregresivo non estacional (AR).
- arima.d: valor numérico. Denota a orde regular de diferenciación.
- arima.q: valor numérico. Denota a orde do polinomio da medias móbiles non estacional (MA).
- arima.bp: valor numérico. Denota a orde do polinomio autorregresivo estacional (AR).
- arima.bd: valor numérico. Denota a orde de diferenciación estacional.
- arima.bq: valor numérico. Denota a orde do polinomio da medias móbiles estacional (MA).

Os nosos modelos, en particular, foron calculados en JDemetra+ e posteriormente fixáronse os parámetros resultantes en R, pero poderiase realizar todo o procedemento dende R. De modo análogo ao visto no Capítulo 3, começariamos por establecer unha selección automática de atípicos e modelo ARIMA e a partir da análise gráfica e dos resultados de validación do modelo, entre outros criterios de interese, iríamos modificando o modelo inicial ata chegar a un novo modelo de axuste estacional cun comportamento razoable.

4.3. Exemplo coa serie do índice de produción industrial de Galicia (IPI15XE)

Unha vez temos as especificacións, é dicir, para cada serie de tempo chamamos á función de especificacións introducíndolle as particularidades da serie, estamos en condicións de realizar o axuste estacional, que se leva a cabo coa metodoloxía TRAMO-SEATS simplemente executando o seguinte liña de código:

```
ts model <- tramoseats(datos.new, spec = especificaciones)
```

onde ‘datos.new’ denota os datos da serie en formato de serie de tempo (ts) e ‘spec’ son as especificacións que lle indicamos previamente, é dicir, o conxunto de parámetros que se introduciron por defecto na función de especificacións (equivalentes aos mostrados na Figura 4.1 para o programa JDemetra+).

Para ver todas estas ideas de forma más clara mostraremos un exemplo a partir dunha serie de tempo das dispoñibles no Banco de series de Galicia. En particular, consideraremos unha das series presentadas no Capítulo 3: o índice xeral de producción industrial de Galicia, na base de 2015 (IPI15XE).

1. Comezamos cargando os datos da serie dende o Banco de series de concxuntura de Galicia no obxecto 'datos_web':

```
datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/IPI15XE")
```

2. A continuación, convertiremos os datos a formato de serie de tempo, indicando en que ano e mes comezan ditos datos. A data de comezo dos datos pódese obter facilmente de multiples formas, como por exemplo imprimindo os primeiros datos da serie mediante ‘head(datos_web)’.

```
datos.new.ts(datos_web[, "dato"], start=c(2002,1), frequency = 12)
```

3. Por outra parte, definimos as especificacións particulares para IPI15XE chamando á función de especificacións, que denotamos previamente co nome *spc*. En particular, nas especificacións temos unha serie de características precedidas do símbolo # para que non as cargue, isto débese a que estas especificacións por defecto xa están establecidas dese modo, polo que será indiferente que o programa as lea ou non.

```
arima.bp = 0, arima.bd=1, arima.bq = 1)
```

En canto aos atípicos e ao modelo ARIMA, introducíense as especificacións do modelo que obtivemos como óptimo no Capítulo 3, e ademais, indicámoslle que non considere o efecto do ano bisiesto, xa que, como o regresor de calendario se introduce nas especificacións en formato ‘usrdef.var’ ou variable definida polo usuario, o test de inclusión ou non do efecto do ano bisiesto non funciona sobre este regresor, polo que teremos que incluílo ou eliminalo de forma manual.

- Neste punto xa estamos en condicións de executar o axuste estacional co modelo TRAMO-SEATS,

```
ts_model <- tramoseats(datos.new, spec = especificaciones)
```

Podemos imprimir os resultados das estimacións dos coeficientes de regresión do modelo accedendo á parte regarima deste (a explicación detallada pódese ver na Sección 2.1.3), ‘ts_model\$regarima’, que nos devolve por pantalla os seguintes resultados:

```
y = regression model + arima (0, 1, 1, 0, 1, 1)
Log-transformation: no
Coefficients:
              Estimate Std. Error
Theta(1)      -0.4381    0.061
BTheta(1)     -0.6662    0.054

              Estimate Std. Error
userdef        0.7547    0.070
Easter [6]     -6.9391   1.104
LS (11-2006)   17.2596   3.885
AO (7-2008)    19.1217   3.664
AO (4-2020)    -39.2506   3.917
AO (5-2020)    -23.2703   3.952

Residual standard error: 4.637 on 222 degrees of freedom
Log likelihood = -685.8, aic = 1390 aicc = 1390, bic(corrected for length) = 3.257
```

Unha vez temos o modelo calculado compre facer a validación deste, así como a posterior representación gráfica da serie corrixida, que, por unha parte pode valer como criterio engadido de validación do modelo e por outra mostra o resultado final do axuste visualmente.

- En canto á validación do modelo temos á nosa disposición todos os test de diagnose presentados no Capítulo 2.

Podemos comezar falando dos test de normalidade, independencia e linearidade dos residuos, que se obteñen como segue:

```
ts_model$regarima$residuals.stat$tests
```

```
Normality           Statistic P.value
mean       -0.57979  0.5626 ***
skewness   -0.02705  0.8667 ***
kurtosis    3.38209  0.2359 **

Signif. codes: H0 (normality of residuals) is not rejected at
significance levels: 0.1 ***0.05 **

Independence          Statistic P.value
ljung box            26.51892  0.2301 ***
ljung box (residuals at seasonal lags) 0.89300  0.6399 ***
```

```

Signif. codes: H0 (independence of residuals) is not rejected at
significance levels: 0.1 ***0.05 **

Linearity
          Statistic P.value
Ljung Box (squared residuals) 18.98528 0.6462 ***
Signif. codes: H0 (no conditional heteroscedasticity of residuals) is not rejected at
significance levels: 0.1 ***0.05 **

```

Á vista dos resultados, podemos apreciar que todos os p-valores son elevados, polo que non temos motivos para rexeitar as hipóteses nulas de normalidade, independencia e linearidade dos residuos, respectivamente. Noutras palabras, os test devolven resultados favorables á normalidade, independencia e linearidade dos residuos, que é o desexable, reforzando a idea de que o modelo obtido é adecuado.

A continuación, accedendo ao obxecto ‘diagnostics’ obtéñense os resultados dos test de presenza de estacionalidade na serie e nos residuos. É dicir, se executamos o comando

```
ts_model$diagnostics
```

obtemos tres bloques de resultados claramente diferenciados,

```

Relative contribution of the components to the stationary
portion of the variance in the original series,
after the removal of the long term trend
Trend computed by Hodrick-Prescott filter (cycle length = 8.0 years)
Component
Cycle      12.829
Seasonal    30.600
Irregular   3.156
TD & Hol.   16.466
Others      44.424
Total       107.475

Combined test in the entire series
Non parametric tests for stable seasonality
                                         P.value
Kruskall-Wallis test                      0.000
Test for the presence of seasonality assuming stability 0.000
Evolutive seasonality test                 0.034

Identifiable seasonality present

Residual seasonality tests
                                         P.value
qs test on sa                            1.000
qs test on i                             1.000
f-test on sa (seasonal dummies)         1.000
f-test on i (seasonal dummies)          1.000
Residual seasonality (entire series)    1.000
Residual seasonality (last 3 years)     0.816
f-test on sa (td)                        0.001
f-test on i (td)                         0.006

```

O primeiro destes bloques mostra a descomposición da varianza total da serie entre as componentes que a forman, máis esta información non será do noso interese. Os dous bloques seguintes serán os de maior interese para a validación do modelo, xa que neles preséntanse os test da presenza de estacionalidade na serie bruta e nos residuos, respectivamente.

No bloque chamado ‘combined test’ temos tres probas sobre a existencia de estacionalidade na serie orixinal. Para que a serie axustada sexa adecuada, estes tests debería apoiar a hipótese de que a serie orixinal presenta estacionalidade. Logo, por construción dos tests, que veñen explicados no Capítulo 2, os dous primeiros deberían ter p-valores pequenos, mostrando probas significativas a

favor da hipótese alternativa (H_1), sendo

$$\begin{aligned} H_0 &: \text{Serie sen estacionalidade,} \\ H_1 &: \text{Serie con estacionalidade,} \end{aligned}$$

mentres que o terceiro test postula as hipóteses ao revés,

$$\begin{aligned} H_0 &: \text{Serie con estacionalidade,} \\ H_1 &: \text{Serie sen estacionalidade,} \end{aligned}$$

polo que neste último o idóneo sería un p-valor grande, indicando que non hai evidencias para rexeitar a hipótese nula (H_0). En particular, os dous primeiros tests si presentan p-valores pequenos, mentres que o terceiro vai máis axustado de cara a ter un p-valor grande pois vale 0.034. A un nivel de significación do 5 %, que é o que se emprega por defecto, rexeitaríase a hipótese nula, polo que debemos tratar con precaución o axuste e considerar os resultados dos restantes test para a validación do modelo, que neste caso son favorables.

Por último todos os test de estacionalidade dos residuos postulan como hipótese nula que os residuos non teñen estacionalidade, fronte á alternativa de que si a teñen. En consecuencia, os p-valores deben ser grandes, indicando que non existen evidencias para rexeitar a hipótese nula. Todos os test teñen, como era de esperar, p-valores grandes, salvo os dous últimos, realizados sobre Working days. Con todo, cabe mencionar que os test realizados sobre Working days ou sobre o efecto de calendario non se poden ter en conta ao traballar en R Project, pois como xa adiantabamos, as variables de calendario foron introducidas como variables definidas polo usuario, de modo que o programa non as identifica completamente como debería para calcular os test, de feito, no Cadro 4.3, podemos apreciar que ditos tests si se comportan ben (devolven p-valores elevados, a favor da hipótese nula) se os executamos en JDemetra+.

residual trading days test	
f-test on sa (td)	Good (0.202)
f-test on i (td)	Good (0.232)

Cadro 4.3: Test de estacionalidade de residuos para Trading days da serie IPI15XE sobre o axustes feitos para o modelo calculado por defecto por JDemetra+ incluíndo as especificacións do usuario.

5. Unha vez temos a diagnose do modelo feita e comprobamos que é adecuado para desestacionalizar a serie de tempo de interese, chega a parte da representación gráfica.

Á hora de representar graficamente, interésanos levar a cabo as mesmas series que hai dispoñibles no Banco de series de concuntura de Galicia, é dicir,

- a serie orixinal,
- a serie correxida de efecto calendario,
- a serie correxida de calendario e estacionalidade,
- a serie correxida de calendario, estacionalidade e atípicos aditivos,
- a serie de tendencia.

Para representar ditas series comezaremos mostrando certa información que temos dispoñible, gardada dentro do obxecto que contén o axuste estacional, ‘ts_model’, información que será necesaria para calcular as series correxidas. Por un lado, temos un elemento chamado ‘final’ que dispón das series e das

predicóns para: a serie orixinal (y), a serie axustada estacionalmente (sa), a serie da compoñente de tendencia (t), a serie da compoñente de estacionalidade (s) e a serie da compoñente irregular (i). Logo, se nos quedamos soamente coas series ('ts_model\$final\$series'), e imprimimos os doce primeiros valores que nos devolve obtense unha táboa coas series de datos dos cales podemos observar os doce primeiros no Cadro 4.4.

	y	sa	t	s	i
2002 Jan	118.58	114.95	115.47	3.63	-0.51
2002 Feb	114.02	114.26	116.09	-0.24	-1.82
2002 Mar	116.73	118.71	117.09	-1.98	1.62
2002 Apr	125.24	117.46	117.95	7.78	-0.49
2002 May	128.94	121.03	118.31	7.90	2.72
2002 Jun	119.69	116.65	118.16	3.04	-1.51
2002 Jul	127.95	117.75	118.10	10.20	-0.35
2002 Aug	94.17	117.59	118.50	-23.41	-0.91
2002 Sep	124.13	119.47	119.18	4.66	0.30
2002 Oct	132.63	120.68	119.76	11.95	0.92
2002 Nov	123.51	118.36	120.31	5.16	-1.96
2002 Dec	112.67	120.87	121.38	-8.20	-0.52

Cadro 4.4: Series finais de IPI15XE.

Ademais, tamén podemos obter as compoñentes deterministas do modelo, coa liña de código

```
ts_model$regarima$model$effects
```

e se imprimimos de novo os doce primeiros valores que nos devolve obtemos os resultados que se presentan no Cadro 4.5.

Polo tanto, temos todos os elementos necesarios para construír as diferentes series que serán obxecto de representación. O procedemento a seguir á hora de xerar as diferentes correccións sobre a serie apoiase nas indicacións presentadas nas notas metodolóxidas do IGE (Instituto Galego de Estatística, 2018). Así, podemos construír as series como segue:

- Serie orixinal: será a propia serie bruta dos datos ('y').
- Serie correxida de efecto calendario: serie orixinal restándolle os efectos deterministas debidos ao calendario, que se amosan no Cadro 4.5: 'tde' (efecto Trading days/Working days), 'ee' (efecto Semana santa) e 'omhe' (outros efectos de calendario). A representación gráfica desta serie xunto coa serie de datos brutos móstrase na Figura 4.2 arriba á esquerda.

	y_lin	tde	ee	omhe	out_t	out_s	out_i	out
2002 Jan	132.45	3.40	0.00	0.00	-17.26	0.00	0.00	-17.26
2002 Feb	131.28	0.00	0.00	0.00	-17.26	0.00	0.00	-17.26
2002 Mar	129.92	7.55	-3.47	0.00	-17.26	0.00	0.00	-17.26
2002 Apr	137.52	1.51	3.47	0.00	-17.26	0.00	0.00	-17.26
2002 May	145.44	0.75	0.00	0.00	-17.26	0.00	0.00	-17.26
2002 Jun	133.18	3.77	0.00	0.00	-17.26	0.00	0.00	-17.26
2002 Jul	143.70	1.51	0.00	0.00	-17.26	0.00	0.00	-17.26
2002 Aug	112.57	-1.13	0.00	0.00	-17.26	0.00	0.00	-17.26
2002 Sep	142.52	-1.13	0.00	0.00	-17.26	0.00	0.00	-17.26
2002 Oct	145.74	4.15	0.00	0.00	-17.26	0.00	0.00	-17.26
2002 Nov	140.77	0.00	0.00	0.00	-17.26	0.00	0.00	-17.26
2002 Dec	129.92	0.00	0.00	0.00	-17.26	0.00	0.00	-17.26

Cadro 4.5: Serie IPI15XE linealizada xunto coas compoñentes determinista.

- Serie corregida de efecto calendario e estacionalidade: será a serie axustada estacionalmente ('sa'), cuxos valores iniciais se mostran no Cadro 4.4 para a serie de IPI15XE. A representación gráfica desta serie xunto coa serie de datos brutos móstrase na Figura 4.2 arriba á dereita.
- Serie corregida de efecto de calendario, estacionalidade e atípicos aditivos: serie 'sa' modificando os valores dos datos atípicos aditivos, restando o valor real ou bruto destes polo valor estimado no modelo para dito atípico (valores que se amosan no Cadro 4.5). A representación gráfica desta serie xunto coa serie de datos brutos móstrase na Figura 4.2 abaixo á esquerda.
- Serie de tendencia: será a serie da compoñente de tendencia como se indica no Cadro 4.4 ('t'). A representación gráfica desta serie xunto coa serie de datos brutos móstrase na Figura 4.2 abaixo á dereita.

Cómpre facer un par de comentarios relacionados coa corrección da serie de atípicos aditivos. O procedemento levado a cabo require do acceso á posición exacta do AO na serie de tempo, de modo que para axilizar esta tarefa empregouse a librería *xts* (Ryan et al., 2020). Ademais, para acceder de forma automática ao valor estimado de cada atípico aditivo, en función da data correspondente, tamén se empregaron as ferramentas dispoñibles na librería *lubridate* (Spinu et al., 2021). Estas librerías axúdanos a construír unha función que realice a corrección de atípicos aditivos sen máis que darlle como datos de entrada: a serie de datos brutos e o modelo calculado con TRAMO-SEATS. Con todo, cabe mencionar que debido a que as series trimestrais teñen unha codificación levemente diferente (trimestres en vez de meses), débese modificar a función de corrección de atípicos para elas, indicándolle que debe detectar as datas en formato trimestre. Vexamos un exemplo a continuación, mostrando como se indican as datas no caso da serie IPI15XE mensual e da serie EPA11_INA trimestral. Podemos apreciar facilmente esta dife-

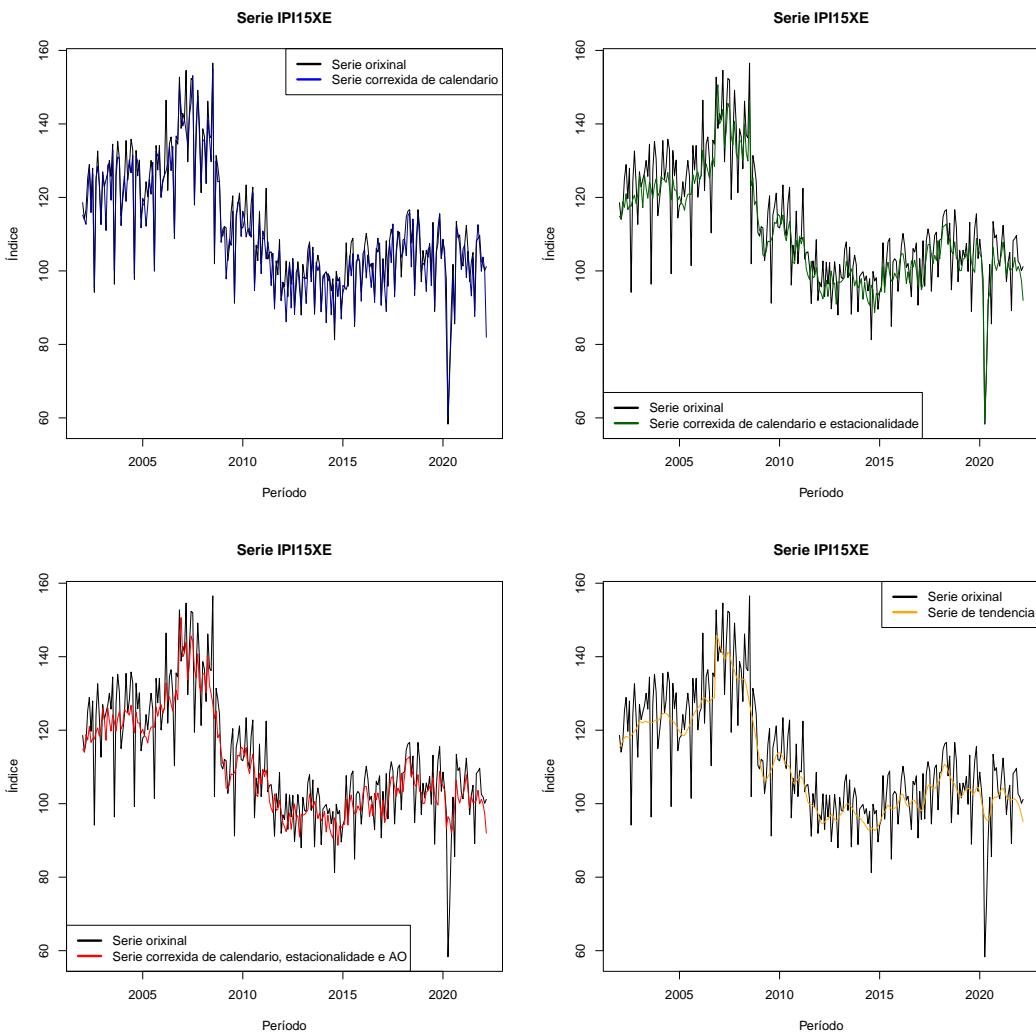


Figura 4.2: Serie orixinal de IPI15XE, xunto coa serie corregida de calendario, en cor azul, arriba á esquerda; xunto coa serie corregida de calendario e estacionalidade, en cor verde, arriba á dereita; xunto coa serie corregida de calendario, estacionalidade e atípicos aditivos, en cor vermella, abaixo á esquerda; xunto coa serie de tendencia, en cor laranxa, abaixo á dereita. Axustes feitos para o modelo calculado por defecto por JDemetra+ incluíndo as modificacións do usuario.

renza se para cada unha das imprimimos os coeficientes de regresión estimados para os valores atípicos, de modo que ao executar a liña de código

```
ts_model$regarima$regression.coefficients
```

obtemos para o IPI15XE os resultados que se mostran no Cadro 4.6, mentres que para a EPA11_INA os resultados preséntanse no formato do Cadro 4.7. Polo tanto, para acceder aos coeficientes estimados, no caso mensual aplicaremos sobre o atípico aditivo a función *month()* e na trimestral a función *quarter()*, convertendo posteriormente o resultado a números romanos (*as.roman()*), de modo que con estas dúas funcións xa podemos corregir todas as series de atípicos aditivos, independentemente de que sexan

	Estimate	Std. Error	T-stat
userdef	0.75	0.07	10.71
Easter [6]	-6.94	1.10	-6.28
LS (11-2006)	17.26	3.89	4.44
AO (7-2008)	19.12	3.66	5.22
AO (4-2020)	-39.25	3.92	-10.02
AO (5-2020)	-23.27	3.95	-5.89

Cadro 4.6: Estimacións dos coeficientes, error típico e estatístico dos efectos deterministas da serie IPI15XE.

	Estimate	Std. Error	T-stat
LS (I-2001)	0.04	0.01	5.10
AO (IV-2004)	-0.01	0.01	-2.30
LS (II-2020)	0.04	0.01	5.34

Cadro 4.7: Estimacións dos coeficientes, error típico e estatístico dos efectos deterministas da serie EPA11_INA.

mensuais ou trimestrais.

Acabamos de ver as ideas intuitivas de como se constrúen as series correxidas para a súa posterior representación gráfica, con todo, a implementación detallada da corrección das series vén recollida no Apéndice B.

4.4. Series na Base de datos

Axustamos 35 series, dispoñibles no Banco de series de convuxura de Galicia, consideradas como as máis relevantes para explicar a convuxura económica de Galicia, co procedemento visto para R Project (adaptado de JDemetra+). As especificacións finais obtidas para un correcto axuste estacional de cada unha destas series de tempo están recollidas no Apéndice A. Entre ditas especificacións podemos distinguir:

- Log: indica se a serie foi transformada logarítmicamente ou non.
- ARIMA: móstranse os valores p, d, q, P, D e Q do modelo ARIMA considerado ($\text{ARIMA}(p, d, q)(P, D, Q)$).
- Media: indica se se considera un modelo ARIMA con media ou sen ela.
- TD/WD: indica se o efecto dos días laborables aparece no modelo como Trading days (6 regresores), como Working days (1 regresor) ou directamente non hai efecto de días laborables.
- LY: indica se se considera efecto de ano bisiesto (Leap Year) ou non.
- Easter: indica se se considera efecto de Semana Santa ou non.

- Outliers: móstranse os outliers que foron incluídos no modelo.

Ademais, as correspondentes series correxidas introducíronse na Base de datos do IGE. Explicaremos a continuación brevemente o proceso levado a cabo para adaptar as series finais obtidas para introducillas na Base de datos do IGE.

Para cargar as series correxidas na Base de datos precisamos que teñan un formato de lectura particular. Os datos de entrada débense presentar nunha táboa cuxas variables sexan: o nome da serie, mes dos datos, ano dos datos e valor que toma a serie nesa data, como se presentan, por exemplo, os datos no Cadro 4.8. Para isto, de novo recorremos as vantaxes da librería *xts*, e construímos unha función chamada *tabla* que xera as táboas de datos para cada serie de interese, das 35 orixinais. É dicir, obtemos 35 táboas, de forma que en cada unha delas teremos os datos da serie dos datos brutos, a serie correxida de calendario, a serie correxida de calendario e estacionalidade, e a serie correxida de calendario, estacionalidade e atípicos aditivos e a serie de tendencia.

	serie	mes	ano	nome
ene. 2002	118.58	1	2002	IPI15XE
feb. 2002	114.02	2	2002	IPI15XE
mar. 2002	116.73	3	2002	IPI15XE
abr. 2002	125.24	4	2002	IPI15XE
may. 2002	128.94	5	2002	IPI15XE
jun. 2002	119.69	6	2002	IPI15XE

Cadro 4.8: Táboa co formato dos datos para a súa lectura na Base de datos.

Por último, unha vez todos os datos están no formato desexado, creamos unha lista que conteña cada unha das táboas para as 35 series, e a continuación sobre esa lista aplicamos a función *do.call()* que terá dous argumentos, o primeiro, ‘rbind’, indícale que queremos que una os datos por filas e o segundo será a propia lista que contén todas as series axustadas, de modo que a función xuntará por filas todos os datos nunha única táboa.

Antes de proceder a cargar os datos na Base de datos cómpre realizar un proceso de comprobación e repasar detidamente os resultados. Neste paso detectamos que as series de turismo durante a aparición do COVID-19 sufrieron un terrible declive, de tal xeito que os seus valores en marzo de 2020 acadaron valores nulos (xa que existía un confinamento vixente). Isto provocou que no axuste realizado sobre elas, en puntos próximos a esta data tómasen valores negativos, o cal non ten sentido. Polo tanto, por coherencia, faremos unha limpeza dos poucos valores negativos existentes nestas series, forzándoo a tomar o valor cero.

O proceso descrito neste capítulo é un mero resumo do procedemento levado a cabo para a modelización das 35 series obxecto de estudo sobre as que se realizou o axuste estacional, que se pode consultar de forma detallada no Apéndice B.

Se comparamos os resultados do Capítulo 3 e os do Capítulo 4, podería parecer que en esencia o proceso é o mesmo. De todos os xeitos, debemos remarcar que o obxectivo neste capítulo é implementar o axuste estacional das series temporais en R mediante un procedemento automático. Isto ten a vantaxe de evitar o procedemento manual que se levou a cabo no Capítulo 3 para a corrección das series. Desta maneira contamos cun programa de R que se executa automaticamente sen pasar constantemente por un procedemento manual cada vez que algunha das series ten que ser actualizada. Aínda así, cada ano haberá que facer unha revisión dos modelos.

Sen embargo, non é o mesmo repetir o proceso manual de cero unha vez ao ano que cada vez que se actualiza algunha das series.

Como desvantaxe, cabe mencionar que os paquetes que dan soporte á adaptación de JDemetra+ para R están aínda en desenvolvemento polo que en R sufrimos pequenas limitacións. Algunhas destas limitacións, xa mencionadas previamente, están relacionadas co calendario de Galicia. Por exemplo, os test de diagnose da estacionalidade dos residuos para o efecto de calendario non se realizan correctamente debido a que o efecto de calendario se incluí no modelo como unha variable externa definida polo usuario, de modo que non se capta o efecto exactamente como debería. Do mesmo xeito, a forma na que se introduce o efecto de calendario no modelo impide que se poida realizar unha proba para decidir se se usan Trading days ou Working days.

Ademais, como é lóxico, JDemetra+ dispón de máis ferramentas das que dispón a adaptación para R, mais ningunha daquelas que están disponibles en JDemetra+ e non en R resulta imprescindible no sentido de que non a precisamos á hora de realizar o axuste estacional correctamente. Ademais, debemos mencionar que se fixo un estudo de comparación pormenorizado dos resultados obtidos en JDemetra+ e R Project comprobando que non hai grandes diferenzas, polo que son procedementos equivalentes a nivel de calidade da estimación do modelo. Polo tanto, aínda que JDemetra+ é un software más completo en canto a funcións e resultados, ás vantaxes que nos supón ter un procedemento automático fanno decantarnos sin lugar a dúbida polo procedemento automático implementado en R, que en particular, é o que queda vixente no IGE para realizar o axuste estacional das serie temporais do Banco de series de conxuntura de Galicia.

Por último para cerrar este capítulo faremos un resumo dos pasos principais seguidos en R para realizar o axuste estacional para unha serie de tempo:

PASO 1: Cargar datos da serie económica de interese.

PASO 2: Convertir datos a formato serie de tempo.

PASO 3: Cargar especificacións por defecto.

PASO 4: Axustar o modelo con TRAMO-SEATS.

PASO 5: Diagnose do modelo e validación gráfica.

- Modelo adecuado: Fin do axuste estacional.
- Modelo inadecuado:
 1. Análise gráfico de posibles atípicos.
 2. Modificar as especificacións por defecto (atípicos e ARIMA).
 3. Volver ao Paso 4.

PASO 6: Calcular as series correxidas de interese (correixa de estacionalidade, de calendario, etc).

PASO 7: Gardar todas as series nunha táboa cun formato que permita a súa lectura na Base de datos do IGE.

Capítulo 5

Conclusíons

Ao longo deste traballo vimos como realizar o axuste estacional das series de conxuntura económica de Galicia seleccionadas. Comezamos introducindo o programa recomendado por Eurostat, JDemetra+, como unha aplicación de axuste estacional de series de tempo que xurde como combinación da metodoloxía TRAMO-SEATS e X-13ARIMA-SEATS, permitindo a opción de empregar calquera dos dous métodos de forma independente. Desenvolveuse a nivel teórico todo o procedemento de axuste estacional implementado internamente en JDemetra+, acompañado da presentación de diferentes test de validación dispoñibles no programa. Ademais, púxose de manifesto que se seguirán as recomendacións de Eurostat de traballar coa metodoloxía TRAMO-SEATS.

Unha vez presentados todos os elementos e mecanismos teóricos presentes detrás de JDemetra+, pasamos ao seguinte paso, que consiste na posta en práctica do axuste estacional. Traballouse co programa JDemetra+, coas series de interese. En termos xerais, pártese dunhas especificacións por defecto establecidas polo programa, considerando o calendario de Galicia entre elas, e en función dos resultados do axuste gráfico e validación do modelo vaise modificando este modelo de partida, que se obtivo case de forma automática a partir do programa. As modificacións realizañse coa idea de non provocar grandes cambios no modelo, no caso de que fose aceptable ou con boas propiedades que desexamos conservar. Outro rasgo a ter en conta é que se buscará a simplicidade do modelo, na medida do posible.

O seguinte paso céntrase na automatización do proceso manual anterior, empregando a linguaxe de programación R. Para pasar todo o procedemento levado a cabo en JDemetra+ a R Project, traballouse coa librería *RJDemetra*, creada precisamente para facilitar unha interface moi próxima a JDemetra+. Con todo cabe mencionar que esta interface está na actualidade en desenvolvemento e tivemos que recorrer a outra librería de apoio, no noso caso para poder considerar o calendario de Galicia. A idea sería introducir en R as especificacións que en JDemetra+ aparecían por defecto e proceder de forma análoga á vista en JDemetra+, é dicir, calcúlase o modelo con estas especificacion. En función dos resultados de validación e do axuste gráfico, realizaremos pequenas modificacións para mellorar o modelo ata chegar a un considerado como óptimo.

Unha vez temos os axustes estacionais realizados en R para as 35 series que foron obxecto de estudio neste traballo, debemos calcular as series correxidas equivalentes ás dispoñibles no Banco de series de conxuntura do IGE, é dicir, as series: orixinal, correxida de efecto calendario, correxida de calendario e estacionalidade, correxida de calendario, estacionalidade e atípicos aditivos e de tendencia. Estas series debén introducirse na Base de datos que nutre o Banco de series de conxuntura, pero antes de proceder coa actualización das series debemos revisalas. Neste paso apreciamos que as series de turismo axustadas na pandemia do COVID-19 presentan valores negativos, debido a que a serie orixinal está moi próxima ou vale directamente cero nese período. Estes valores negativos mostran incoherencia coa vida real, polo que son substituídos por cero. Este paso queda pendente de revisión, xa que no futuro deberase establecer un procedemento que estude e corrixa este tipo de situacións de forma automática. Ademais, para introducir as series na Base de datos debemos renomear certas variables, para adecualas ao formato de lectura que está pre establecido na Base de datos.

Chegados a este punto podemos dar conta dunha serie de melloras acadadas para o axuste estacional das series de conjuntura económica de Galicia. En primeiro lugar, a aparición do COVID-19 en marzo de 2020 tivo terribles consecuencias para as series axustadas que estaban vixentes naquel momento, xa que os datos a partir de marzo de 2020 mostraron comportamentos inestables, con picos e cambios de tendencia, que tiveron que ser corregidos mediante diferentes tipos de outliers, segundo as recomendacións de Eurostat. Unha vez se introduciron os atípicos para correxir os citados comportamentos anómalos, acádanse de novo series cun comportamento desexable.

Ademais, todo o procedemento de axuste estacional levouse a cabo en JDemetra+ coa posterior adaptación en R Project, polo que finalmente obtivemos un procedemento automático para R que axusta as series de interese. Polo tanto, as series finais obtidas teñen un axuste estacional satisfactorio que non precisa ser modificado a curto prazo. Con todo, como obxectivo futuro, cómpre facer unha revisión do axuste estacional, para analizar posibles fructuacións que poidan dar lugar a unha necesidade de cambios no modelo.

Outra mellora que queda aberta á súa posta en práctica no futuro, será a utilización das prediccións que temos dispoñibles xunto coas series axustadas e que permitirían dispoñer dunha ferramenta moi útil para o seguimento da conjuntura económica de Galicia.

Por último, o proceso automático implementado en R queda na actualidade en funcionamento no IGE para o axuste estacional das series económicas, e pódese consultar no Banco de series de conjuntura (Instituto Galego de Estatística, 2022). Nel inclúense os axustes estacionais froito do traballo levado a cabo nesta memoria, acompañados dun manual auxiliar con indicacións sobre a metodoloxía e o proceso levado a cabo en R para obter o axuste estacional das series de interese.

Apéndice A

Táboa de parámetros dos modelos finais

A continuación preséntase a táboa coas especificacións finais que dan forma aos modelos de axuste estacional construídos para cada unha das 35 series obxecto de estudio neste traballo. En particular a primeira columna da táboa contén o nome da serie e a continuación temos unha columna con cada especificación. As especificacións indícanos: a estructura do modelo final (Log, ARIMA e Media), os efectos de calendario considerados (TD/WD, LY e Easter) e os atípicos incluídos no modelo (Outliers). De forma máis precisa,

- Log: indica se se considera a transformación logarítmica sobre a serie de datos brutos ou non.
- ARIMA: indica a orde da parte regular e da parte estacional para o modelo ARIMA.
- Media: indica se se considera o modelo ARIMA con media ou sen ela.
- TD/WD: indica se o efecto dos días laborables se inclúe no modelo ou non, e en caso de que se inclúa, se se consideran regresores de tipo Trading days (TD) ou Working days (WD).
- LY: indica se se inclúe o efecto de ano bisiesto (Leap Year).
- Easter: indica se se inclúe o efecto de Semana Santa.
- Outliers: indica se se consideran atípicos ou non, e en caso afirmativo, móstrase o tipo de atípico e a data na que se produce.

Cadro A.1: Modelos finais.

Serie	Log	ARIMA	Media	TD/WD	LY	Easter	Outliers
AMSS	Si	(1,1,1)(0,1,1)	Non	Non	Non	Non	AO 3/2020 LS 4/2020
PAROREG	Si	(1,1,1)(1,1,1)	Non	Non	Non	Non	AO 3/2020 AO 4/2020
EPA11_OCUT	Si	(3,1,0)(0,1,1)	Non	Non	Non	Non	LS II/2020

Continua na páxina seguinte

Cadro A.1 – Continuación da páxina anterior

Serie	Log	ARIMA	Media	TD/WD	LY	Easter	Outliers
EPA11_PART	Non	(0,1,1)(0,1,1)	Non	Non	Non	Non	LS II/2020
EPA11_INA	Si	(3,1,1)(0,1,1)	Non	Non	Non	Non	LS I/2001 AO IV/2004 LS II/2020
IPI15XE	Non	(0,1,1)(0,1,1)	Non	WD	Non	Si	LS 11/2006 AO 7/2008 AO 4/2020 AO 5/2020
PRD_VEH	Non	(3,0,0)(0,1,1)	Non	WD	Non	Si	AO 3/2020 LS 4/2020 LS 6/2020 LS 5/2021
CPPGASOLT	Non	(0,1,1)(0,1,1)	Non	TD	Si	Si	AO 3/2020 LS 4/2020
CPPGO	Si	(0,1,1)(0,1,1)	Non	WD	Si	Si	Non
GEETO	Si	(1,0,0)(0,1,1)	Non	Non	Non	Non	Non
ENRTO	Si	(0,1,0)(1,0,0)	Non	Non	Non	Non	Non
ERTO	Si	(0,1,1)(0,1,1)	Non	Non	Non	Non	Non
VIINI	Si	(0,1,1)(0,1,1)	Non	Non	Non	Non	LS 6/1980 AO 8/1980 LS 1/1998
VITER	Si	(0,1,1)(1,1,1)	Non	Non	Non	Non	LS 1/1996 LS 8/2008

Continua na páxina seguinte

Cadro A.1 – Continuación da páxina anterior

Serie	Log	ARIMA	Media	TD/WD	LY	Easter	Outliers
							TC 8/2012 AO 1/2016 AO 7/2017 AO 2/2020 LS 3/2020
EDTVIV	Si	(0,1,1)(1,1,1)	Non	Non	Non	Non	AO 1/2015
EDTSUP	Si	(0,1,1)(1,1,1)	Non	Non	Non	Non	Non
SERV_CN_2015	Si	(2,1,0)(0,1,1)	Non	WD	Si	Si	LS 2/2020 LS 6/2020 LS 2/2021 LS 11/2021
SERV_OCUP_2015	Si	(0,1,1)(0,1,1)	Non	Non	Non	Non	LS 3/2020 AO 5/2020 LS 6/2020
TAP	Non	(3,1,1)(0,1,1)	Non	Non	Non	Si	AO 9/2014 LS 3/2020 LS 8/2020
TMM	Si	(0,1,2)(0,1,1)	Non	Non	Non	Non	AO 4/1987 AO 2/1989
EOHP_04	Non	(0,1,1)(0,1,1)	Non	Non	Non	Si	LS 3/2020 LS 5/2020
EOHVE_04	Non	(0,1,0)(0,1,1)	Non	Non	Non	Si	LS 3/2020 LS 5/2020

Continua na páxina seguinte

Cadro A.1 – Continuación da páxina anterior

Serie	Log	ARIMA	Media	TD/WD	LY	Easter	Outliers
EOHGR_04	Non	(0,1,1)(0,1,1)	Non	Non	Non	Si	LS 3/2020 LS 5/2020
EOTRP_04	Non	(1,0,0)(1,1,1)	Non	WD	Non	Si	LS 3/2020 LS 5/2020
EOTRVE_04	Non	(1,0,0)(0,1,1)	Non	WD	Non	Si	LS 3/2020 LS 5/2020 LS 7/2021
EOTRGR_04	Non	(0,1,1)(0,1,1)	Non	WD	Non	Si	LS 5/2020
CPM_15_XERAL_09	Non	(0,1,1)(0,1,1)	Non	TD	Si	Non	AO 4/2020 AO 5/2020
IPRI15XE	Si	(0,2,2)(0,1,1)	Non	Non	Non	Non	LS 5/2020 AO 1/2021 LS 9/2021
MATTUR	Non	(0,1,1)(0,1,1)	Non	WD	Si	Si	AO 3/2020
MATVCARGA	Non	(0,1,1)(0,1,1)	Non	WD	Non	Si	AO 12/2007 AO 3/2020
CXM	Si	(0,1,1)(0,1,1)	Non	WD	Si	Si	AO 4/2020 AO 5/2020
CXX	Si	(0,1,1)(0,1,1)	Non	WD	Non	Non	AO 4/2020 AO 5/2020 AO 10/2021
IVU10_MTOT	Non	(1,1,0)(0,1,1)	Non	Non	Non	Non	Non
IVU10_XTOT	Si	(2,1,0)(0,1,1)	Non	Non	Non	Non	Non

Continua na páxina seguinte

Cadro A.1 – Continuación da páxina anterior

Serie	Log	ARIMA	Media	TD/WD	LY	Easter	Outliers
IVU10_MNE	Non	(0,1,1)(0,1,1)	Non	Non	Non	Non	AO 11/1999 LS 9/2000

Apéndice B

Código R

Neste apartado inclúese o código completo de R que permite realizar o axuste estacional das series de tempo automaticamente. Podemos estructurar o código nas seguintes partes:

1. Cargar librerías.
2. Construir calendario de Galicia.
3. Construir as dúas opcións de regresores de calendario: Working days e Trading days.
4. Implementar función de especificacións para o axuste con TRAMO-SEATS.
5. Implementar función coa serie correxida de estacionalidade, calendario e atípicos aditivos para
 - a) series mensuais,
 - b) series trimestrais.
6. Implementar función táboa co conxunto de series correxidas xunto coa orixinal, correspondentes a cada serie de datos brutos considerada.
7. Axuste estacional de cada unha das 35 series obxecto de estudo e obtención das series correxidas.
8. Crear unha táboa coas 35 series orixinais e as correspondentes series correxidas para cada unha delas.

```
## Cargar librerías:  
library(rjd3modelling)  
library(RJDemetra)  
library(lubridate)  
library(xts)  
library(car)  
  
## Calendario Galicia:  
GaliciaCalendar <- calendar.new()  
calendar.holiday(GaliciaCalendar, "NEWYEAR")  
calendar.fixedday(GaliciaCalendar, month = 1, day = 6)  
calendar.holiday(GaliciaCalendar, "MAYDAY")  
calendar.fixedday(GaliciaCalendar, month = 5, day = 17)  
calendar.fixedday(GaliciaCalendar, month = 7, day = 25)  
calendar.holiday(GaliciaCalendar, "ASSUMPTION")  
calendar.fixedday(GaliciaCalendar, month = 10, day = 12)  
calendar.holiday(GaliciaCalendar, "ALLSAINTSDAY")  
calendar.fixedday(GaliciaCalendar, month = 12, day = 6)
```



```

# - Función especificaciones JDemetra+: #

spc <- function(outlier.enabled = FALSE,
                  usrdef.outliersEnabled = TRUE,
                  usrdef.outliersType, usrdef.outliersDate,
                  usrdef.varEnabled = TRUE, usrdef.var=regresseurs_JO,
                  easter.type = "IncludeEaster",
                  automdl.enabled = FALSE,
                  arima.mu= FALSE,
                  arima.p = 0, arima.d=1, arima.q = 1,
                  arima.bp = 0, arima.bd=1, arima.bq = 1 ){

  tramoseats_spec(spec = "RSAfull",
                   preliminary.check = TRUE,
                   estimate.tol = 0.0000001,
                   estimate.eml = TRUE,
                   estimate.urfinal = 0.96,
                   transform.function = "Auto",
                   transform.fct = 0.95,
                   usrdef.outliersEnabled = usrdef.outliersEnabled,
                   usrdef.outliersType = usrdef.outliersType,
                   usrdef.outliersDate = usrdef.outliersDate,
                   usrdef.varEnabled = usrdef.varEnabled,
                   usrdef.var = usrdef.var,
                   usrdef.varType = "Calendar",
                   tradingdays.option = "UserDefined",
                   tradingdays.mauto = "FTest",
                   tradingdays.pftd = 0.01,
                   # tradingdays.leapyear = TRUE,
                   # tradingdays.test = "Separate_T",
                   easter.type = easter.type,
                   easter.julian = FALSE,
                   easter.duration = 6,
                   easter.test = TRUE,
                   outlier.enabled = outlier.enabled,
                   outlier.usedefcv = TRUE,
                   outlier.cv = 3.5,
                   outlier.ao = TRUE,
                   outlier.tc = TRUE,
                   outlier.ls = TRUE,
                   outlier.so = FALSE,
                   outlier.eml = FALSE,
                   outlier.tcrate = 0.7,
                   automdl.enabled = automdl.enabled,
                   automdl.acceptdefault = FALSE,
                   automdl.cancel = 0.05,
                   automdl.ub1 = 0.97,
                   automdl.ub2 = 0.91,
                   automdl.armalimit = 1,
                   automdl.reducecv = 0.12,
                   automdl.ljungboxlimit = 0.95,
                   automdl.compare = FALSE,
                   arima.mu= arima.mu,

```

```

arima.p = arima.p, arima.d=arima.d, arima.q = arima.q,
arima.bp = arima.bp, arima.bd=arima.bd, arima.bq = arima.bq
,
arima.coefEnabled = TRUE,
seats.predictionLength = -1,
seats.approx = "Legacy",
seats.maBoundary = 0.95,
seats.trendBoundary = 0.5,
seats.seasTol = 2,
seats.seasdBoundary = 0.8,
seats.seasdBoundary1 = 0.8,
seats.method = "Burman")
}

# - Función Serie correxida de estacionalidade, calendario e AO #

### Mensual

s_correx_S_C_AO_mensual<-function (serie,modelo){
  ZZA_m<-as.xts(modelo$final$series[, "sa"])
  ano1<-start(serie)[1]
  mes1<-start(serie)[2]
  atipicos<-modelo$regarima$specification$regression$userdef$outliers
  if (is.na(atipicos)){
    print("Non hai atípicos")
  } else {
    dat_atip<-atipicos[atipicos$type=="AO",]$date
    n<-length(dat_atip)
    anoAO<-numeric(n)
    mesAO<-numeric(n)
    est_AO<-numeric(n)
    if(n==0) {
      print("Non hai atípicos aditivos (AO)")
    } else {
      for (i in 1:n){
        anoAO[i]<-year(dat_atip[i])
        mesAO[i]<-month(dat_atip[i])
        est_AO[i]<-modelo$regarima$regression.coefficients
          [sprintf("AO (%s-%s)", mesAO[i],anoAO[i]),1]
        ZZA_m[sprintf("%s %s",anoAO[i],mesAO[i]),]<-ZZA_m[
          sprintf("%s %s",anoAO[i],mesAO[i]),]-est_AO[i]
      }
    }
  }
  ZZA_m<-ts(ZZA_m,start=start(ZZA_m),end = end(ZZA_m),frequency=12)
  return(ZZA_m)
}

### Trimestral

s_correx_S_C_AO_trimestral<-function (serie,modelo){
  ZZA_t<-as.xts(modelo$final$series[, "sa"])
  ano1<-start(serie)[1]

```



```

usrdef.outliersDate = c("2020-03-01", "2020-04-01"),
#usrdef.varEnabled = TRUE,
#usrdef.var = regresseurs_J0,
#easter.type = "IncludeEaster",
#outlier.enabled = FALSE,
#automdl.enabled = FALSE,
#arima.mu= FALSE,
arima.p = 1, arima.d=1, arima.q = 1,
arima.bp = 0, arima.bd=1, arima.bq = 1)

ts_model <- tramoseats(datos.new, spec = especificaciones) # TRAMO-SEATS method

## Como extraer coeficientes modelo: ##
ts_model$regarima

## Como obter diagnose modelo: ##
ts_model$diagnostics
ts_model$regarima$residuals.stat$tests

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
AMSS<-series_datos[, "y"]

### Serie correxida de calendario:
AMSSZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
deterministas[, "omhe"])

### Serie correxida de calendario e estacionalidade:
AMSSZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
AMSSZZA<-s_correx_S_C_AO_mensual(datos.new,ts_model)

### Serie de tendencia:
AMSSZT<-series_datos[, "t"]

final<-cbind(AMSS,AMSSZZC,AMSSZZS,AMSSZZA,AMSSZT)
resultados_AMSS<-tabla(final)
resultados_AMSS$nome<-Recode(resultados_AMSS$nome,"'AMSSZZC'='AMSS&ZZC';
'AMSSZZS'='AMSS&ZZS'; 'AMSSZZA'='AMSS&ZZA';
'AMSSZT'='AMSS&ZT'" )

#####
#### PAROREG #####
#####

datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/PAROREG")
datos.new=ts(datos_web[, "dato"], start=c(2001,2), frequency = 12)

especificaciones<-spc(#usrdef.outliersEnabled = TRUE,

```

```

usrdef.outliersType = c("AO", "AO"),
usrdef.outliersDate = c("2020-03-01", "2020-04-01"),
#usrdef.varEnabled = TRUE,
#usrdef.var = regresseurs_JO,
#easter.type = "IncludeEaster",
#outlier.enabled = FALSE,
#automdl.enabled = FALSE,
#arima.mu= FALSE,
arima.p = 1, arima.d=1, arima.q = 1,
arima.bp = 1, arima.bd=1, arima.bq = 1)

ts_model <- tramoseats(datos.new, spec = especificaciones)

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
PAROREG<-series_datos[, "y"]

### Serie correxida de calendario:
PAROREGZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
deterministas[, "omhe"])

### Serie correxida de calendario e estacionalidade:
PAROREGZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
PAROREGZZA<-s_correx_S_C_AO_mensual(datos.new,ts_model)

### Serie de tendencia:
PAROREGZT<-series_datos[, "t"]

final<-cbind(PAROREG, PAROREGZZC, PAROREGZZS, PAROREGZZA, PAROREGZT)
resultados_PAROREG<-tabla(final)
resultados_PAROREG$nome<-Recode(resultados_PAROREG$nome,
                                     "'PAROREGZZC'='PAROREG&ZZC',
                                     'PAROREGZZS'='PAROREG&ZZS',
                                     'PAROREGZZA'='PAROREG&ZZA',
                                     'PAROREGZT'='PAROREG&ZT'" )

#####
#### EPA11_OCUT #####
#####

datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/EPA11_OCUT")
datos.new=ts(datos_web[, "dato"], start=c(1996,1), frequency = 4)

especificaciones<-spc(#usrdef.outliersEnabled = TRUE,
                      usrdef.outliersType = c("LS"),
                      usrdef.outliersDate = c("2020-04-01"),
                      #usrdef.varEnabled = TRUE,
                      #usrdef.var = regresseurs_JO,

```

```

#easter.type = "IncludeEaster",
#outlier.enabled = FALSE,
#automdl.enabled = FALSE,
#arima.mu= FALSE,
arima.p = 3, arima.d=1, arima.q = 0,
arima.bp = 0, arima.bd=1, arima.bq = 1)

ts_model <- tramoseats(datos.new, spec = especificaciones)

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
EPA11_OCUT<-series_datos[, "y"]

### Serie correxida de calendario:
EPA11_OCUTZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
deterministas[, "omhe"])

### Serie correxida de calendario e estacionalidade:
EPA11_OCUTZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
EPA11_OCUTZZA<-s_correx_S_C_AO_trimestral(datos.new,ts_model)

### Serie de tendencia:
EPA11_OCUTZZT<-series_datos[, "t"]

final<-cbind(EPA11_OCUT,EPA11_OCUTZZC,EPA11_OCUTZZS,EPA11_OCUTZZA,EPA11_OCUTZZT)
resultados_EPA11_OCUT<-tabla(final)
resultados_EPA11_OCUT$mes<-Recode(resultados_EPA11_OCUT$mes, "1=13; 4=14; 7=15;
10=16")
resultados_EPA11_OCUT$nome<-Recode(resultados_EPA11_OCUT$nome,
"'EPA11_OCUTZZC'='EPA11_OCUT&ZZC';
'EPA11_OCUTZZS'='EPA11_OCUT&ZZS';
'EPA11_OCUTZZA'='EPA11_OCUT&ZZA';
'EPA11_OCUTZZT'='EPA11_OCUT&ZZT' " )

#####
#### EPA11_PART ####
#####

datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/EPA11_PART")
datos.new=ts(datos_web[, "dato"],start=c(1996,1),frequency = 4)

especificaciones<-spc(#usrdef.outliersEnabled = TRUE,
usrdef.outliersType = c("LS"),
usrdef.outliersDate = c("2020-04-01"),
#usrdef.varEnabled = TRUE,
#usrdef.var = regresseurs_JO,
#easter.type = "IncludeEaster",

```

```

#outlier.enabled = FALSE,
#automdl.enabled = FALSE,
#arima.mu= FALSE,
arima.p = 0, arima.d=1, arima.q = 1,
arima.bp = 0, arima.bd=1, arima.bq = 1)

ts_model <- tramoseats(datos.new, spec = especificaciones)

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
EPA11_PART<-series_datos[, "y"]

### Serie correxida de calendario:
EPA11_PARTZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
deterministas[, "omhe"])

### Serie correxida de calendario e estacionalidade:
EPA11_PARTZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
EPA11_PARTZZA<-s_correx_S_C_AO_trimestral(datos.new,ts_model)

### Serie de tendencia:
EPA11_PARTZZT<-series_datos[, "t"]

final<-cbind(EPA11_PART,EPA11_PARTZZC,EPA11_PARTZZS,EPA11_PARTZZA,EPA11_PARTZZT)
resultados_EPA11_PART<-tabla(final)
resultados_EPA11_PART$mes<-Recode(resultados_EPA11_PART$mes, "1=13; 4=14; 7=15;
10=16")
resultados_EPA11_PART$nome<-Recode(resultados_EPA11_PART$nome,
"‘EPA11_PARTZZC’='EPA11_PART&ZZC';
‘EPA11_PARTZZS’='EPA11_PART&ZZS';
‘EPA11_PARTZZA’='EPA11_PART&ZZA';
‘EPA11_PARTZZT’='EPA11_PART&ZZT' " )

#####
#### EPA11_INA #####
#####

datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/EPA11_INA")
datos.new=ts(datos_web[, "dato"],start=c(1996,1),frequency = 4)

especificaciones<-spc(#usrdef.outliersEnabled = TRUE,
usrdef.outliersType = c("LS", "AO", "LS"),
usrdef.outliersDate = c("2001-01-01", "2004-10-01", "
2020-04-01"),
usrdef.varEnabled = TRUE,
usrdef.var = regresseurs_JO,
easter.type = "Unused",

```

```

#outlier.enabled = FALSE,
#automdl.enabled = FALSE,
#arima.mu= FALSE,
arima.p = 3, arima.d=1, arima.q = 1,
arima.bp = 0, arima.bd=1, arima.bq = 1)

ts_model <- tramoseats(datos.new, spec = especificaciones)

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
EPA11_INA<-series_datos[,"y"]

### Serie correxida de calendario:
EPA11_INAZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
deterministas[, "omhe"])

### Serie correxida de calendario e estacionalidade:
EPA11_INAZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
EPA11_INAZZA<-s_correx_S_C_AO_trimestral(datos.new,ts_model)

### Serie de tendencia:
EPA11_INAZZT<-series_datos[, "t"]

final<-cbind(EPA11_INA,EPA11_INAZZC,EPA11_INAZZS,EPA11_INAZZA,EPA11_INAZZT)
resultados_EPA11_INA<-tabla(final)
resultados_EPA11_INA$mes<-Recode(resultados_EPA11_INA$mes, "1=13; 4=14; 7=15;
10=16")
resultados_EPA11_INA$nome<-Recode(resultados_EPA11_INA$nome,
                                     "'EPA11_INAZZC'='EPA11_INA&ZZC';
'EPA11_INAZZS'='EPA11_INA&ZZS';
'EPA11_INAZZA'='EPA11_INA&ZZA';
'EPA11_INAZZT'='EPA11_INA&ZT'")

#####
#### IPI15XE #####
#####

datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/IPI15XE")
datos.new=ts(datos_web[, "dato"], start=c(2002,1), frequency = 12)

especificaciones<-spc(#usrdef.outliersEnabled = TRUE,
                       usrdef.outliersType = c("LS", "AO", "AO", "AO"),
                       usrdef.outliersDate = c("2006-11-01", "2008-07-01", "
2020-04-01","2020-05-01"),
                       #usrdef.varEnabled = TRUE,
                       usrdef.var = regresseurs_JO[, "week_Days"],

```

```

#easter.type = "IncludeEaster",
#outlier.enabled = FALSE,
#automdl.enabled = FALSE,
#arima.mu= FALSE,
arima.p = 0, arima.d=1, arima.q = 1,
arima.bp = 0, arima.bd=1, arima.bq = 1)

ts_model <- tramoseats(datos.new, spec = especificaciones)

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
IPI15XE<-series_datos[, "y"]

### Serie correxida de calendario:
IPI15XEZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
deterministas[, "omhe"])

### Serie correxida de calendario e estacionalidade:
IPI15XEZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
IPI15XEZZA<-s_correx_S_C_AO_mensual(datos.new,ts_model)

### Serie de tendencia:
IPI15XEZZT<-series_datos[, "t"]

final<-cbind(IPI15XE, IPI15XEZZC, IPI15XEZZS, IPI15XEZZA, IPI15XEZZT)
resultados_IPI15XE<-tabla(final)
resultados_IPI15XE$nome<-Recode(resultados_IPI15XE$nome,
                                    "'IPI15XEZZC'='IPI15XE&ZZC';
                                    'IPI15XEZZS'='IPI15XE&ZS';
                                    'IPI15XEZZA'='IPI15XE&ZZA';
                                    'IPI15XEZZT'='IPI15XE&ZT'" )

#####
#### PRD_VEH #####
#####

datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/PRD_VEH")
datos.new=ts(datos_web[, "dato"], start=c(2003,1), frequency = 12)

especificaciones<-spc(#usrdef.outliersEnabled = TRUE,
                      usrdef.outliersType = c("AO", "LS", "LS", "LS"),
                      usrdef.outliersDate = c("2020-03-01", "2020-04-01", "
2020-06-01", "2021-05-01"),
                      #usrdef.varEnabled = TRUE,
                      usrdef.var = regresseurs_JO[, "week_Days"],
                      #easter.type = "IncludeEaster",
                      #outlier.enabled = FALSE,
                      #automdl.enabled = FALSE,

```

```

#arima.mu= FALSE,
arima.p = 3, arima.d=0, arima.q = 0,
arima.bp = 0, arima.bd=1, arima.bq = 1)

ts_model <- tramoseats(datos.new, spec = especificaciones)

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
PRD_VEH<-series_datos[, "y"]

### Serie correxida de calendario:
PRD_VEHZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
deterministas[, "omhe"])

### Serie correxida de calendario e estacionalidade:
PRD_VEHZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
PRD_VEHZZA<-s_correx_S_C_AO_mensual(datos.new,ts_model)

### Serie de tendencia:
PRD_VEHZZT<-series_datos[, "t"]

final<-cbind(PRD_VEH, PRD_VEHZZC, PRD_VEHZZS, PRD_VEHZZA, PRD_VEHZZT)
resultados_PRD_VEH<-tabla(final)
resultados_PRD_VEH$nome<-Recode(resultados_PRD_VEH$nome,
                                     "'PRD_VEHZZC'='PRD_VEH&ZZC';
                                     'PRD_VEHZZS'='PRD_VEH&ZS';
                                     'PRD_VEHZZA'='PRD_VEH&ZZA';
                                     'PRD_VEHZZT'='PRD_VEH&ZT'" )

#####
#### CPPGASOLT #####
#####

datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/CPPGASOLT")
datos.new=ts(datos_web[, "dato"], start=c(1993,1), frequency = 12)

especificaciones<-spc(#usrdef.outliersEnabled = TRUE,
                      usrdef.outliersType = c("AO", "LS"),
                      usrdef.outliersDate = c("2020-03-01", "2020-04-01"),
                      #usrdef.varEnabled = TRUE,
                      usrdef.var = regresseurs_JO_TD,
                      #easter.type = "IncludeEaster",
                      #outlier.enabled = FALSE,
                      #automdl.enabled = FALSE,
                      #arima.mu= FALSE,
                      arima.p = 0, arima.d=1, arima.q = 1,
                      arima.bp = 0, arima.bd=1, arima.bq = 1)

```

```

ts_model <- tramoseats(datos.new, spec = especificaciones)

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
CPPGASOLT<-series_datos[, "y"]

### Serie correxida de calendario:
CPPGASOLTZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
    deterministas[, "omhe"])

### Serie correxida de calendario e estacionalidade:
CPPGASOLTZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
CPPGASOLTZZA<-s_correx_S_C_AO_mensual(datos.new,ts_model)

### Serie de tendencia:
CPPGASOLTZZT<-series_datos[, "t"]

final<-cbind(CPPGASOLT,CPPGASOLTZZC,CPPGASOLTZZS,CPPGASOLTZZA,CPPGASOLTZZT)
resultados_CPPGASOLT<-tabla(final)
resultados_CPPGASOLT$nome<-Recode(resultados_CPPGASOLT$nome,
    "'CPPGASOLTZZC'='CPPGASOLT&ZZC';
    'CPPGASOLTZZS'='CPPGASOLT&ZZS';
    'CPPGASOLTZZA'='CPPGASOLT&ZZA';
    'CPPGASOLTZZT'='CPPGASOLT&ZZT'" )

#####
#### CPPGO #####
#####

datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/CPPGO")
datos.new=ts(datos_web[, "dato"], start=c(1993,1), frequency = 12)

especificaciones<-spc(#usrdef.outliersEnabled = TRUE,
    usrdef.outliersType = c(NA),
    usrdef.outliersDate = c(NA),
    #usrdef.varEnabled = TRUE,
    #usrdef.var = regresseurs_JO,
    #easter.type = "IncludeEaster",
    #outlier.enabled = FALSE,
    #automdl.enabled = FALSE,
    #arima.mu= FALSE,
    arima.p = 0, arima.d=1, arima.q = 1,
    arima.bp = 0, arima.bd=1, arima.bq = 1)

ts_model <- tramoseats(datos.new, spec = especificaciones)

## Series correxidas ##

```

```

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
CPPGO<-series_datos[, "y"]

### Serie correxida de calendario:
CPPGOZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
deterministas[, "omhe"])

### Serie correxida de calendario e estacionalidade:
CPPGOZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
CPPGOZZA<-s_correx_S_C_AO_mensual(datos.new, ts_model)

### Serie de tendencia:
CPPGOZT<-series_datos[, "t"]

final<-cbind(CPPGO, CPPGOZZC, CPPGOZZS, CPPGOZZA, CPPGOZT)
resultados_CPPGO<-tabla(final)
resultados_CPPGO$nome<-Recode(resultados_CPPGO$nome, "'CPPGOZZC'='CPPGO&ZZC';
'CPPGOZZS'='CPPGO&ZZS'; 'CPPGOZZA'='CPPGO&ZZA';
'CPPGOZT'='CPPGO&ZT' ")

#####
#### GEETO #####
#####

datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/GEETO")
datos.new=ts(datos_web[, "dato"], start=c(2013,1), frequency = 12)

specificacions<-spc(#usrdef.outliersEnabled = TRUE,
                      usrdef.outliersType = c(NA),
                      usrdef.outliersDate = c(NA),
                      #usrdef.varEnabled = TRUE,
                      #usrdef.var = regresseurs_J0,
                      #easter.type = "IncludeEaster",
                      #outlier.enabled = FALSE,
                      #automdl.enabled = FALSE,
                      #arima.mu= FALSE,
                      arima.p = 1, arima.d=0, arima.q = 0,
                      arima.bp = 0, arima.bd=1, arima.bq = 1)

ts_model <- tramoseats(datos.new, spec = specificacions)

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
GEETO<-series_datos[, "y"]

```

```

### Serie correxida de calendario:
GEETOZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
deterministas[, "omhe"])

### Serie correxida de calendario e estacionalidade:
GEETOZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
GEETOZZA<-s_correx_S_C_AO_mensual(datos.new,ts_model)

### Serie de tendencia:
GEETOZZT<-series_datos[, "t"]

final<-cbind(GEETO,GEETOZZC,GEETOZZS,GEETOZZA,GEETOZZT)
resultados_GEETO<-tabla(final)
resultados_GEETO$nome<-Recode(resultados_GEETO$nome,"'GEETOZZC'='GEETO&ZZC';
'GEETOZZS'='GEETO&ZZS'; 'GEETOZZA'='GEETO&ZZA';
'GEETOZZT'='GEETO&ZTT'" )

#####
#### ENRTO #####
#####

datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/ENRTO")
datos.new=ts(datos_web[, "dato"], start=c(2013,1), frequency = 12)

specificacions<-spc(#usrdef.outliersEnabled = TRUE,
                      usrdef.outliersType = c(NA),
                      usrdef.outliersDate = c(NA),
                      #usrdef.varEnabled = TRUE,
                      #usrdef.var = regresseurs_J0,
                      #easter.type = "IncludeEaster",
                      #outlier.enabled = FALSE,
                      #automdl.enabled = FALSE,
                      #arima.mu= FALSE,
                      arima.p = 0, arima.d=1, arima.q = 0,
                      arima.bp = 1, arima.bd=0, arima.bq = 0)

ts_model <- tramoseats(datos.new, spec = specificacions)

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
ENRTO<-series_datos[, "y"]

### Serie correxida de calendario:
ENRTOZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
deterministas[, "omhe"])

```

```

### Serie correxida de calendario e estacionalidade:
ENRTOZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
ENRTOZZA<-s_correx_S_C_AO_mensual(datos.new, ts_model)

### Serie de tendencia:
ENRTOZT<-series_datos[, "t"]

final<-cbind(ENRTO, ENRTOZZC, ENRTOZZS, ENRTOZZA, ENRTOZT)
resultados_ENRTO<-tabla(final)
resultados_ENRTO$nome<-Recode(resultados_ENRTO$nome, "'ENRTOZZC'='ENRTO&ZZC';
                                'ENRTOZZS'='ENRTO&ZZS'; 'ENRTOZZA'='ENRTO&ZZA';
                                'ENRTOZT'='ENRTO&ZT' " )

#####
#### ERTO ####
#####

datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/ERTO")
datos.new=ts(datos_web[, "dato"], start=c(2013,1), frequency = 12)

specificacions<-spc(#usrdef.outliersEnabled = TRUE,
                      usrdef.outliersType = c(NA),
                      usrdef.outliersDate = c(NA),
                      #usrdef.varEnabled = TRUE,
                      #usrdef.var = regresseurs_J0,
                      #easter.type = "IncludeEaster",
                      #outlier.enabled = FALSE,
                      #automdl.enabled = FALSE,
                      #arima.mu= FALSE,
                      arima.p = 0, arima.d=1, arima.q = 1,
                      arima.bp = 0, arima.bd=1, arima.bq = 1)

ts_model <- tramoseats(datos.new, spec = specificacions)

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
ERTO<-series_datos[, "y"]

### Serie correxida de calendario:
ERTOZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
    deterministas[, "omhe"])

### Serie correxida de calendario e estacionalidade:
ERTOZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
ERTOZZA<-s_correx_S_C_AO_mensual(datos.new, ts_model)

```

```

### Serie de tendencia:
ERTOZZT<-series_datos[, "t"]

final<-cbind(ERTO, ERTOZZC, ERTOZZS, ERTOZZA, ERTOZZT)
resultados_ERTO<-tabla(final)
resultados_ERTO$nome<-Recode(resultados_ERTO$nome, "'ERTOZZC'='ERTO&ZZC',
                                'ERTOZZS'='ERTO&ZZS'; 'ERTOZZA'='ERTO&ZZA';
                                'ERTOZZT'='ERTO&ZZT'" )

#####
#### VIINI #####
#####

datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/VIINI")
datos.new=ts(datos_web[, "dato"], start=c(1980,1), frequency = 12)

specificacions<-spc(#usrdef.outliersEnabled = TRUE,
                      usrdef.outliersType = c("LS", "AO", "LS"),
                      usrdef.outliersDate = c("1980-06-01", "1980-08-01", "1998-01-01"),
                      #usrdef.varEnabled = TRUE,
                      #usrdef.var = regresseurs_J0,
                      #easter.type = "IncludeEaster",
                      #outlier.enabled = FALSE,
                      #automdl.enabled = FALSE,
                      #arima.mu= FALSE,
                      arima.p = 0, arima.d=1, arima.q = 1,
                      arima.bp = 0, arima.bd=1, arima.bq = 1)

ts_model <- tramoseats(datos.new, spec = specificacions)

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
VIINI<-series_datos[, "y"]

### Serie correxida de calendario:
VIINIZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
    deterministas[, "omhe"])

### Serie correxida de calendario e estacionalidade:
VIINIZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
VIINIZZA<-s_correx_S_C_AO_mensual(datos.new, ts_model)

### Serie de tendencia:
VIINIZZT<-series_datos[, "t"]

final<-cbind(VIINI, VIINIZZC, VIINIZZS, VIINIZZA, VIINIZZT)
resultados_VIINI<-tabla(final)

```



```

#####
##### EDTVIV #####
#####

datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/EDTVIV")
datos.new=ts(datos_web[, "dato"], start=c(1990,1), frequency = 12)

specificacions<-spc(#usrdef.outliersEnabled = TRUE,
                      usrdef.outliersType = c("AO"),
                      usrdef.outliersDate = c("2015-01-01"),
                      #usrdef.varEnabled = TRUE,
                      #usrdef.var = regresseurs_JO,
                      #easter.type = "IncludeEaster",
                      #outlier.enabled = FALSE,
                      #automdl.enabled = FALSE,
                      #arima.mu= FALSE,
                      arima.p = 0, arima.d=1, arima.q = 1,
                      arima.bp = 1, arima.bd=1, arima.bq = 1)

ts_model <- tramoseats(datos.new, spec = specificacions)

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
EDTVIV<-series_datos[, "y"]

### Serie correxida de calendario:
EDTVIVZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
deterministas[, "omhe"])

### Serie correxida de calendario e estacionalidade:
EDTVIVZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
EDTVIVZZA<-s_correx_S_C_AO_mensual(datos.new, ts_model)

### Serie de tendencia:
EDTVIVZZT<-series_datos[, "t"]

final<-cbind(EDTVIV,EDTVIVZZC,EDTVIVZZS,EDTVIVZZA,EDTVIVZZT)
resultados_EDTVIV<-tabla(final)
resultados_EDTVIV$nome<-Recode(resultados_EDTVIV$nome,
                                "'EDTVIVZZC'='EDTVIV&ZZC';
                                'EDTVIVZZS'='EDTVIV&ZZS';
                                'EDTVIVZZA'='EDTVIV&ZZA';
                                'EDTVIVZZT'='EDTVIV&ZZT'" )

#####
##### EDTSUP #####
#####

```

```

#####
datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/EDTSUP")
datos.new=ts(datos_web[, "dato"], start=c(1990,1), frequency = 12)

specificacions<-spc(#usrdef.outliersEnabled = TRUE,
                      usrdef.outliersType = c(NA),
                      usrdef.outliersDate = c(NA),
                      #usrdef.varEnabled = TRUE,
                      #usrdef.var = regresseurs_JO,
                      #easter.type = "IncludeEaster",
                      #outlier.enabled = FALSE,
                      #automdl.enabled = FALSE,
                      #arima.mu= FALSE,
                      arima.p = 0, arima.d=1, arima.q = 1,
                      arima.bp = 1, arima.bd=1, arima.bq = 1)

ts_model <- tramoseats(datos.new, spec = specificacions)

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
EDTSUP<-series_datos[, "y"]

### Serie correxida de calendario:
EDTSUPZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
deterministas[, "omhe"])

### Serie correxida de calendario e estacionalidade:
EDTSUPZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
EDTSUPZZA<-s_correx_S_C_AO_mensual(datos.new, ts_model)

### Serie de tendencia:
EDTSUPZT<-series_datos[, "t"]

final<-cbind(EDTSUP,EDTSUPZZC,EDTSUPZZS,EDTSUPZZA,EDTSUPZT)
resultados_EDTSUP<-tabla(final)
resultados_EDTSUP$nome<-Recode(resultados_EDTSUP$nome,
                                "'EDTSUPZZC'='EDTSUP&ZZC';
                                'EDTSUPZZS'='EDTSUP&ZZS';
                                'EDTSUPZZA'='EDTSUP&ZZA';
                                'EDTSUPZT'='EDTSUP&ZT'" )

#####
##### SERV_CN_2015 #####
#####

datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/SERV_CN_2015")
datos.new=ts(datos_web[, "dato"], start=c(2005,1), frequency = 12)
```

```

specificaciones<-spc(#usrdef.outliersEnabled = TRUE,
                      usrdef.outliersType = c("LS", "LS", "LS", "LS"),
                      usrdef.outliersDate = c("2020-02-01", "2020-06-01", "
                                             2021-02-01","2021-11-01"),
                      #usrdef.varEnabled = TRUE,
                      #usrdef.var = regresseurs_J0,
                      #easter.type = "IncludeEaster",
                      #outlier.enabled = FALSE,
                      #automdl.enabled = FALSE,
                      #arima.mu= FALSE,
                      arima.p = 2, arima.d=1, arima.q = 0,
                      arima.bp = 0, arima.bd=1, arima.bq = 1)

ts_model <- tramoseats(datos.new, spec = specificaciones)

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
SERV_CN_2015<-series_datos[, "y"]

### Serie correxida de calendario:
SERV_CN_2015ZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
  deterministas[, "omhe"])

### Serie correxida de calendario e estacionalidade:
SERV_CN_2015ZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
SERV_CN_2015ZZA<-s_correx_S_C_AO_mensual(datos.new,ts_model)

### Serie de tendencia:
SERV_CN_2015ZZT<-series_datos[, "t"]

final<-cbind(SERV_CN_2015,SERV_CN_2015ZZC,SERV_CN_2015ZZS,SERV_CN_2015ZZA,SERV_CN_
  2015ZZT)
resultados_SERV_CN_2015<-tabla(final)
resultados_SERV_CN_2015$nome<-Recode(resultados_SERV_CN_2015$nome,
                                         "'SERV_CN_2015ZZC'='SERV_CN_2015&ZZC',
                                         'SERV_CN_2015ZZS'='SERV_CN_2015&ZZS',
                                         'SERV_CN_2015ZZA'='SERV_CN_2015&ZZA',
                                         'SERV_CN_2015ZZT'='SERV_CN_2015&ZZT'" )

#####
#### SERV_OCUP_2015 ####
#####

datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/SERV_OCUP_2015"
  )
datos.new=ts(datos_web[, "dato"], start=c(2005,1), frequency = 12)

```

```

specificaciones<-spc(#usrdef.outliersEnabled = TRUE,
                      usrdef.outliersType = c("LS", "AO", "LS"),
                      usrdef.outliersDate = c("2020-03-01", "2020-05-01","2020-06-01
                      "),
                      #usrdef.varEnabled = TRUE,
                      #usrdef.var = regresseurs_JO,
                      #easter.type = "IncludeEaster",
                      #outlier.enabled = FALSE,
                      #automdl.enabled = FALSE,
                      #arima.mu= FALSE,
                      arima.p = 0, arima.d=1, arima.q = 1,
                      arima.bp = 0, arima.bd=1, arima.bq = 1)

ts_model <- tramoseats(datos.new, spec = especificaciones)

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
SERV_OCUP_2015<-series_datos[, "y"]

### Serie correxida de calendario:
SERV_OCUP_2015ZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
deterministas[, "omhe"])

### Serie correxida de calendario e estacionalidade:
SERV_OCUP_2015ZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
SERV_OCUP_2015ZZA<-s_correx_S_C_AO_mensual(datos.new,ts_model)

### Serie de tendencia:
SERV_OCUP_2015ZZT<-series_datos[, "t"]

final<-cbind(SERV_OCUP_2015,SERV_OCUP_2015ZZC,SERV_OCUP_2015ZZS,SERV_OCUP_2015ZZA,
              SERV_OCUP_2015ZZT)
resultados_SERV_OCUP_2015<-tabla(final)
resultados_SERV_OCUP_2015$nome<-Recode(resultados_SERV_OCUP_2015$nome,
                                         "'SERV_OCUP_2015ZZC'='SERV_OCUP_2015&ZZC';
                                         'SERV_OCUP_2015ZZS'='SERV_OCUP_2015&ZZS';
                                         'SERV_OCUP_2015ZZA'='SERV_OCUP_2015&ZZA';
                                         'SERV_OCUP_2015ZZT'='SERV_OCUP_2015&ZZT' " )

#####
#####
### TAP #####
#####

datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/TAP")
datos.new=ts(datos_web[, "dato"], start=c(1980,1), frequency = 12)

specificaciones<-spc(#usrdef.outliersEnabled = TRUE,
                      usrdef.outliersType = c( "AO", "LS", "LS"),

```

```

usrdef.outliersDate = c("2014-09-01", "2020-03-01", "2020-08-01
"),
#usrdef.varEnabled = TRUE,
#usrdef.var = regresseurs_J0,
#easter.type = "IncludeEaster",
#outlier.enabled = FALSE,
#automdl.enabled = FALSE,
#arima.mu= FALSE,
arima.p = 3, arima.d=1, arima.q = 1,
arima.bp = 0, arima.bd=1, arima.bq = 1)

ts_model <- tramoseats(datos.new, spec = especificaciones)

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
TAP<-series_datos[, "y"]
### Serie correxida de calendario:
TAPZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
deterministas[, "omhe"])

### Serie correxida de calendario e estacionalidade:
TAPZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
TAPZZA<-s_correx_S_C_AO_mensual(datos.new,ts_model)

### Serie de tendencia:
TAPZZT<-series_datos[, "t"]

final<-cbind(TAP,TAPZZC,TAPZZS,TAPZZA,TAPZZT)
resultados_TAP<-tabla(final)
resultados_TAP$nome<-Recode(resultados_TAP$nome,"'TAPZZC'='TAP&ZZC';
' TAPZZS='TAP&ZZS'; 'TAPZZA='TAP&ZZA';
' TAPZZT='TAP&ZZT'" )

#####
### TMM #####
#####

datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/TMM")
datos.new=ts(datos_web[, "dato"], start=c(1980,1), frequency = 12)

especificaciones<-spc(#usrdef.outliersEnabled = TRUE,
usrdef.outliersType = c("AO", "AO"),
usrdef.outliersDate = c("1987-04-01", "1989-02-01"),
#usrdef.varEnabled = TRUE,
#usrdef.var = regresseurs_J0,
#easter.type = "IncludeEaster",
#outlier.enabled = FALSE,
#automdl.enabled = FALSE,

```

```

#arima.mu= FALSE,
arima.p = 0, arima.d=1, arima.q = 2,
arima.bp = 0, arima.bd=1, arima.bq = 1)

ts_model <- tramoseats(datos.new, spec = especificaciones)

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
TMM<-series_datos[, "y"]

### Serie correxida de calendario:
TMMZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
deterministas[, "omhe"])

### Serie correxida de calendario e estacionalidade:
TMMZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
TMMZZA<-s_correx_S_C_AO_mensual(datos.new,ts_model)

### Serie de tendencia:
TMMZZT<-series_datos[, "t"]

final<-cbind(TMM,TMMZZC,TMMZZS,TMMZZA,TMMZZT)
resultados_TMM<-tabla(final)
resultados_TMM$nome<-Recode(resultados_TMM$nome,"'TMMZZC'='TMM&ZZC';
'TMMZZS'='TMM&ZZS'; 'TMMZZA'='TMM&ZZA';
'TMMZZT'='TMM&ZZT'")

#####
#### EOHP_04 ####
#####

datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/EOHP_04")
datos.new=ts(datos_web[, "dato"], start=c(1999,1), frequency = 12)

especificaciones<-spc(#usrdef.outliersEnabled = TRUE,
                       usrdef.outliersType = c("LS", "LS"),
                       usrdef.outliersDate = c("2020-03-01", "2020-05-01"),
                       #usrdef.varEnabled = TRUE,
                       #usrdef.var = regresseurs_J0,
                       #easter.type = "IncludeEaster",
                       #outlier.enabled = FALSE,
                       #automdl.enabled = FALSE,
                       #arima.mu= FALSE,
                       arima.p = 0, arima.d=1, arima.q = 1,
                       arima.bp = 0, arima.bd=1, arima.bq = 1)

ts_model <- tramoseats(datos.new, spec = especificaciones)

```

```

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
EOHP_04<-series_datos[, "y"]

### Serie correxida de calendario:
EOHP_04ZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
deterministas[, "omhe"])

### Serie correxida de calendario e estacionalidade:
EOHP_04ZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
EOHP_04ZZA<-s_correx_S_C_AO_mensual(datos.new,ts_model)

### Serie de tendencia:
EOHP_04ZZT<-series_datos[, "t"]

final<-cbind(EOHP_04,EOHP_04ZZC,EOHP_04ZZS,EOHP_04ZZA,EOHP_04ZZT)
resultados_EOHP_04<-tabla(final)
resultados_EOHP_04$nome<-Recode(resultados_EOHP_04$nome,
                                     "'EOHP_04ZZC'='EOHP_04&ZZC';
                                     'EOHP_04ZZS'='EOHP_04&ZZS';
                                     'EOHP_04ZZA'='EOHP_04&ZZA';
                                     'EOHP_04ZZT'='EOHP_04&ZZT'" )

#####
#### EOHVE_04 ####
#####

datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/EOHVE_04")
datos.new=ts(datos_web[, "dato"], start=c(1999,1), frequency = 12)

specificaciones<-spc(#usrdef.outliersEnabled = TRUE,
                      usrdef.outliersType = c("LS", "LS"),
                      usrdef.outliersDate = c("2020-03-01", "2020-05-01"),
                      #usrdef.varEnabled = TRUE,
                      #usrdef.var = regresseurs_J0,
                      #easter.type = "IncludeEaster",
                      #outlier.enabled = FALSE,
                      #automdl.enabled = FALSE,
                      #arima.mu= FALSE,
                      arima.p = 0, arima.d=1, arima.q = 0,
                      arima.bp = 0, arima.bd=1, arima.bq = 1)

ts_model <- tramoseats(datos.new, spec = specificaciones)

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

```

```

### Serie de datos brutos:
EOHVE_04<-series_datos[, "y"]

### Serie correxida de calendario:
EOHVE_04ZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
deterministas[, "omhe"])

### Serie correxida de calendario e estacionalidade:
EOHVE_04ZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
EOHVE_04ZZA<-s_correx_S_C_AO_mensual(datos.new,ts_model)

### Serie de tendencia:
EOHVE_04ZZT<-series_datos[, "t"]

final<-cbind(EOHVE_04, EOHVE_04ZZC, EOHVE_04ZZS, EOHVE_04ZZA, EOHVE_04ZZT)
resultados_EOHVE_04<-tabla(final)
resultados_EOHVE_04$nome<-Recode(resultados_EOHVE_04$nome,
                                     "'EOHVE_04ZZC'='EOHVE_04&ZZC';
                                     'EOHVE_04ZZS'='EOHVE_04&ZZS';
                                     'EOHVE_04ZZA'='EOHVE_04&ZZA';
                                     'EOHVE_04ZZT'='EOHVE_04&ZZT'" )

#####
#### EOHGR_04 #####
#####

datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/EOHGR_04")
datos.new=ts(datos_web[, "dato"], start=c(2004,1), frequency = 12)

specificacions<-spc(#usrdef.outliersEnabled = TRUE,
                      usrdef.outliersType = c("LS", "LS"),
                      usrdef.outliersDate = c("2020-03-01", "2020-05-01"),
                      #usrdef.varEnabled = TRUE,
                      #usrdef.var = regresseurs_J0,
                      #easter.type = "IncludeEaster",
                      #outlier.enabled = FALSE,
                      #automdl.enabled = FALSE,
                      #arima.mu= FALSE,
                      arima.p = 0, arima.d=1, arima.q = 1,
                      arima.bp = 0, arima.bd=1, arima.bq = 1)

ts_model <- tramoseats(datos.new, spec = specificacions)

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
EOHGR_04<-series_datos[, "y"]

```

```

### Serie correxida de calendario:
EOHGR_04ZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
deterministas[, "omhe"])

### Serie correxida de calendario e estacionalidade:
EOHGR_04ZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
EOHGR_04ZZA<-s_correx_S_C_AO_mensual(datos.new,ts_model)

### Serie de tendencia:
EOHGR_04ZZT<-series_datos[, "t"]

final<-cbind(EOHGR_04,EOHGR_04ZZC,EOHGR_04ZZS,EOHGR_04ZZA,EOHGR_04ZZT)
resultados_EOHGR_04<-tabla(final)
resultados_EOHGR_04$nome<-Recode(resultados_EOHGR_04$nome,
                                     "'EOHGR_04ZZC'='EOHGR_04&ZZC';
                                     'EOHGR_04ZZS'='EOHGR_04&ZZS';
                                     'EOHGR_04ZZA'='EOHGR_04&ZZA';
                                     'EOHGR_04ZZT'='EOHGR_04&ZZT'" )

#####
#### EOTRP_04 #####
#####

datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/EOTRP_04")
datos.new=ts(datos_web[, "dato"], start=c(2004,1), frequency = 12)

specificacions<-spc(#usrdef.outliersEnabled = TRUE,
                      usrdef.outliersType = c("LS", "LS"),
                      usrdef.outliersDate = c("2020-03-01", "2020-05-01"),
                      #usrdef.varEnabled = TRUE,
                      usrdef.var = regresseurs_JO[, "week_Days"],
                      #easter.type = "IncludeEaster",
                      #outlier.enabled = FALSE,
                      #automdl.enabled = FALSE,
                      #arima.mu= FALSE,
                      arima.p = 1, arima.d=0, arima.q = 0,
                      arima.bp = 1, arima.bd=1, arima.bq = 1)

ts_model <- tramoseats(datos.new, spec = specificacions)

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
EOTRP_04<-series_datos[, "y"]

### Serie correxida de calendario:
EOTRP_04ZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
deterministas[, "omhe"])

```

```

### Serie correxida de calendario e estacionalidade:
EOTRP_04ZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
EOTRP_04ZZA<-s_correx_S_C_AO_mensual(datos.new, ts_model)

### Serie de tendencia:
EOTRP_04ZZT<-series_datos[, "t"]

final<-cbind(EOTRP_04, EOTRP_04ZZC, EOTRP_04ZZS, EOTRP_04ZZA, EOTRP_04ZZT)
resultados_EOTRP_04<-tabla(final)
resultados_EOTRP_04$nome<-Recode(resultados_EOTRP_04$nome,
                                     "'EOTRP_04ZZC'='EOTRP_04&ZZC',
                                     'EOTRP_04ZZS'='EOTRP_04&ZZS',
                                     'EOTRP_04ZZA'='EOTRP_04&ZZA',
                                     'EOTRP_04ZZT'='EOTRP_04&ZZT'")

#####
#### EOTRVE_04 #####
#####

datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/EOTRVE_04")
datos.new=ts(datos_web[, "dato"], start=c(2004,1), frequency = 12)

specificacions<-spc(#usrdef.outliersEnabled = TRUE,
                      usrdef.outliersType = c("LS", "LS", "LS"),
                      usrdef.outliersDate = c("2020-03-01", "2020-05-01", "2021-07-01"),
                      ),
                      #usrdef.varEnabled = TRUE,
                      usrdef.var = regresseurs_J0[, "week_Days"],
                      #easter.type = "IncludeEaster",
                      #outlier.enabled = FALSE,
                      #automdl.enabled = FALSE,
                      #arima.mu= FALSE,
                      arima.p = 1, arima.d=0, arima.q = 0,
                      arima.bp = 0, arima.bd=1, arima.bq = 1)

ts_model <- tramoseats(datos.new, spec = specificacions)

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
EOTRVE_04<-series_datos[, "y"]

### Serie correxida de calendario:
EOTRVE_04ZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
deterministas[, "omhe"])

### Serie correxida de calendario e estacionalidade:
EOTRVE_04ZZS<-series_datos[, "sa"]

```

```

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
EOTRVE_04ZZA<-s_correx_S_C_AO_mensual(datos.new,ts_model)

### Serie de tendencia:
EOTRVE_04ZZT<-series_datos[, "t"]

final<-cbind(EOTRVE_04,EOTRVE_04ZZC,EOTRVE_04ZZS,EOTRVE_04ZZA,EOTRVE_04ZZT)
resultados_EOTRVE_04<-tabla(final)
resultados_EOTRVE_04$nome<-Recode(resultados_EOTRVE_04$nome,
                                      "'EOTRVE_04ZZC'='EOTRVE_04&ZZC';
                                       'EOTRVE_04ZZS'='EOTRVE_04&ZZS';
                                       'EOTRVE_04ZZA'='EOTRVE_04&ZZA';
                                       'EOTRVE_04ZZT'='EOTRVE_04&ZZT' "
                                     )

#####
#### EOTRGR_04 #####
#####

datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/EOTRGR_04")
datos.new=ts(datos_web[, "dato"], start=c(2004,1), frequency = 12)

specificacions<-spc(#usrdef.outliersEnabled = TRUE,
                      usrdef.outliersType = c("LS"),
                      usrdef.outliersDate = c("2020-05-01"),
                      #usrdef.varEnabled = TRUE,
                      usrdef.var = regresseurs_JO[, "week_Days"],
                      #easter.type = "IncludeEaster",
                      #outlier.enabled = FALSE,
                      #automdl.enabled = FALSE,
                      #arima.mu= FALSE,
                      arima.p = 0, arima.d=1, arima.q = 1,
                      arima.bp = 0, arima.bd=1, arima.bq = 1)

ts_model <- tramoseats(datos.new, spec = specificacions)

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
EOTRGR_04<-series_datos[, "y"]

### Serie correxida de calendario:
EOTRGR_04ZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
  deterministas[, "omhe"])

### Serie correxida de calendario e estacionalidade:
EOTRGR_04ZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
EOTRGR_04ZZA<-s_correx_S_C_AO_mensual(datos.new,ts_model)

### Serie de tendencia:

```

```

EOTRGR_04ZZT<-series_datos[, "t"]

final<-cbind(EOTRGR_04,EOTRGR_04ZZC,EOTRGR_04ZZS,EOTRGR_04ZZA,EOTRGR_04ZZT)
resultados_EOTRGR_04<-tabla(final)
resultados_EOTRGR_04$nome<-Recode(resultados_EOTRGR_04$nome,
                                      "'EOTRGR_04ZZC'='EOTRGR_04&ZZC';
                                       'EOTRGR_04ZZS'='EOTRGR_04&ZZS';
                                       'EOTRGR_04ZZA'='EOTRGR_04&ZZA';
                                       'EOTRGR_04ZZT'='EOTRGR_04&ZZT'" )

#####
##### CPM_15_XERAL_09 #####
#####

datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/CPM_15_XERAL_09
")
datos.new=ts(datos_web[, "dato"], start=c(2006,1), frequency = 12)

specificacions<-spc(#usrdef.outliersEnabled = TRUE,
                      usrdef.outliersType = c("AO", "AO"),
                      usrdef.outliersDate = c("2020-04-01", "2020-05-01"),
                      #usrdef.varEnabled = TRUE,
                      usrdef.var = regresseurs_JO_TD,
                      #easter.type = "IncludeEaster",
                      #outlier.enabled = FALSE,
                      #automdl.enabled = FALSE,
                      #arima.mu= FALSE,
                      arima.p = 0, arima.d=1, arima.q = 1,
                      arima.bp = 0, arima.bd=1, arima.bq = 1)

ts_model <- tramoseats(datos.new, spec = specificacions)

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
CPM_15_XERAL_09<-series_datos[, "y"]

### Serie correxida de calendario:
CPM_15_XERAL_09ZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"
    ]+deterministas[, "omhe"])

### Serie correxida de calendario e estacionalidade:
CPM_15_XERAL_09ZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
CPM_15_XERAL_09ZZA<-s_correx_S_C_AO_mensual(datos.new,ts_model)

### Serie de tendencia:
CPM_15_XERAL_09ZZT<-series_datos[, "t"]

final<-cbind(CPM_15_XERAL_09,CPM_15_XERAL_09ZZC,CPM_15_XERAL_09ZZS,CPM_15_XERAL_09

```

```

ZZA,CPM_15_XERAL_09ZZT)
resultados_CPM_15_XERAL_09<-tabla(final)
resultados_CPM_15_XERAL_09$nome<-Recode(resultados_CPM_15_XERAL_09$nome,
                                         "'CPM_15_XERAL_09ZZC'='CPM_15_XERAL_09&ZZC';
                                         'CPM_15_XERAL_09ZZS'='CPM_15_XERAL_09&ZZS';
                                         'CPM_15_XERAL_09ZZA'='CPM_15_XERAL_09&ZZA';
                                         'CPM_15_XERAL_09ZZT'='CPM_15_XERAL_09&ZZT' " )"

#####
##### IPRI15XE #####
#####

datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/IPRI15XE")
datos.new=ts(datos_web[, "dato"], start=c(2002,1), frequency = 12)

specificacions<-spc(#usrdef.outliersEnabled = TRUE,
                      usrdef.outliersType = c("LS", "AO", "LS"),
                      usrdef.outliersDate = c("2020-05-01", "2021-01-01", "2021-09-01
                                             "),
                      #usrdef.varEnabled = TRUE,
                      #usrdef.var = regresseurs_JO,
                      #easter.type = "IncludeEaster",
                      #outlier.enabled = FALSE,
                      #automdl.enabled = FALSE,
                      #arima.mu= FALSE,
                      arima.p = 0, arima.d=2, arima.q = 2,
                      arima.bp = 0, arima.bd=1, arima.bq = 1)

ts_model <- tramoseats(datos.new, spec = specificacions)

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
IPRI15XE<-series_datos[, "y"]

### Serie correxida de calendario:
IPRI15XEZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
  deterministas[, "omhe"])

### Serie correxida de calendario e estacionalidade:
IPRI15XEZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
IPRI15XEZZA<-s_correx_S_C_AO_mensual(datos.new,ts_model)

### Serie de tendencia:
IPRI15XEZZT<-series_datos[, "t"]

final<-cbind(IPRI15XE,IPRI15XEZZC,IPRI15XEZZS,IPRI15XEZZA,IPRI15XEZZT)
resultados_IPRI15XE<-tabla(final)
resultados_IPRI15XE$nome<-Recode(resultados_IPRI15XE$nome,

```



```

#####
##### MATVCARGA #####
#####

datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/MATVCARGA")
datos.new=ts(datos_web[, "dato"], start=c(1985,1), frequency = 12)

specificacions<-spc(#usrdef.outliersEnabled = TRUE,
                      usrdef.outliersType = c("AO", "AO"),
                      usrdef.outliersDate = c("2007-12-01", "2020-03-01"),
                      #usrdef.varEnabled = TRUE,
                      usrdef.var = regresseurs_JO[, "week_Days"],
                      #easter.type = "IncludeEaster",
                      #outlier.enabled = FALSE,
                      #automdl.enabled = FALSE,
                      #arima.mu= FALSE,
                      arima.p = 0, arima.d=1, arima.q = 1,
                      arima.bp = 0, arima.bd=1, arima.bq = 1)

ts_model <- tramoseats(datos.new, spec = specificacions)

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
MATVCARGA<-series_datos[, "y"]

### Serie correxida de calendario:
MATVCARGAZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
  deterministas[, "omhe"])

### Serie correxida de calendario e estacionalidade:
MATVCARGAZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
MATVCARGAZZA<-s_correx_S_C_AO_mensual(datos.new,ts_model)

### Serie de tendencia:
MATVCARGAZZT<-series_datos[, "t"]

final<-cbind(MATVCARGA,MATVCARGAZZC,MATVCARGAZZS,MATVCARGAZZA,MATVCARGAZZT)
resultados_MATVCARGA<-tabla(final)
resultados_MATVCARGA$nome<-Recode(resultados_MATVCARGA$nome,
                                      "'MATVCARGAZZC'='MATVCARGA&ZZC';
                                       'MATVCARGAZZS'='MATVCARGA&ZZS';
                                       'MATVCARGAZZA'='MATVCARGA&ZZA';
                                       'MATVCARGAZZT'='MATVCARGA&ZZT' " )
```

```

##### CXM #####
#####
#datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/CXM")
#datos.new=ts(datos_web[, "dato"], start=c(1988,1), frequency = 12)

specificacions<-spc(#usrdef.outliersEnabled = TRUE,
                      usrdef.outliersType = c("AO", "AO"),
                      usrdef.outliersDate = c("2020-04-01", "2020-05-01"),
                      #usrdef.varEnabled = TRUE,
                      #usrdef.var = regresseurs_J0,
                      #easter.type = "IncludeEaster",
                      #outlier.enabled = FALSE,
                      #automdl.enabled = FALSE,
                      #arima.mu= FALSE,
                      arima.p = 0, arima.d=1, arima.q = 1,
                      arima.bp = 0, arima.bd=1, arima.bq = 1)

ts_model <- tramoseats(datos.new, spec = specificacions)

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
CXM<-series_datos[, "y"]

### Serie correxida de calendario:
CXMZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
  deterministas[, "omhe"])

### Serie correxida de calendario e estacionalidade:
CXMZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
CXMZZA<-s_correx_S_C_AO_mensual(datos.new, ts_model)

### Serie de tendencia:
CXMZZT<-series_datos[, "t"]

final<-cbind(CXM, CXMZZC, CXMZZS, CXMZZA, CXMZZT)
resultados_CXM<-tabla(final)
resultados_CXM$nome<-Recode(resultados_CXM$nome, "'CXMZZC'='CXM&ZZC';
                                'CXMZZS'='CXM&ZZS'; 'CXMZZA'='CXM&ZZA';
                                'CXMZZT'='CXM&ZZT'")

#####
##### CXX #####
#####

#datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/CXX")
#datos.new=ts(datos_web[, "dato"], start=c(1988,1), frequency = 12)

```

```

specificaciones<-spc(#usrdef.outliersEnabled = TRUE,
                      usrdef.outliersType = c("AO", "AO", "AO"),
                      usrdef.outliersDate = c("2020-04-01", "2020-05-01", "2021-10-01
                      "),
                      #usrdef.varEnabled = TRUE,
                      usrdef.var = regresseurs_JO[, "week_Days"],
                      #easter.type = "IncludeEaster",
                      #outlier.enabled = FALSE,
                      #automdl.enabled = FALSE,
                      #arima.mu= FALSE,
                      arima.p = 0, arima.d=1, arima.q = 1,
                      arima.bp = 0, arima.bd=1, arima.bq = 1)

ts_model <- tramoseats(datos.new, spec = specificaciones)

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
CXX<-series_datos[, "y"]

### Serie correxida de calendario:
CXXZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
deterministas[, "omhe"])

### Serie correxida de calendario e estacionalidade:
CXXZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
CXXZZA<-s_correx_S_C_AO_mensual(datos.new,ts_model)

### Serie de tendencia:
CXXZT<-series_datos[, "t"]

final<-cbind(CXX,CXXZZC,CXXZZS,CXXZZA,CXXZT)
resultados_CXX<-tabla(final)
resultados_CXX$nome<-Recode(resultados_CXX$nome, "'CXXZZC'='CXX&ZZC';
'CXXZZS'='CXX&ZZS'; 'CXXZZA'='CXX&ZZA';
'CXXZT'='CXX&ZT' " )

#####
#### IVU10_MTOT #####
#####

datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/IVU10_MTOT")
datos.new=ts(datos_web[, "dato"], start=c(1995,1), frequency = 12)

specificaciones<-spc(#usrdef.outliersEnabled = TRUE,
                      usrdef.outliersType = c(NA),
                      usrdef.outliersDate = c(NA),
                      #usrdef.varEnabled = TRUE,
                      #usrdef.var = regresseurs_JO,

```

```

#easter.type = "IncludeEaster",
#outlier.enabled = FALSE,
#automdl.enabled = FALSE,
#arima.mu= FALSE,
arima.p = 1, arima.d=1, arima.q = 0,
arima.bp = 0, arima.bd=1, arima.bq = 1)

ts_model <- tramoseats(datos.new, spec = especificaciones)

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
IVU10_MTOT<-series_datos[, "y"]

### Serie correxida de calendario:
IVU10_MTOTZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
deterministas[, "omhe"])

### Serie correxida de calendario e estacionalidade:
IVU10_MTOTZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
IVU10_MTOTZZA<-s_correx_S_C_AO_mensual(datos.new, ts_model)

### Serie de tendencia:
IVU10_MTOTZZT<-series_datos[, "t"]

final<-cbind(IVU10_MTOT, IVU10_MTOTZZC, IVU10_MTOTZZS, IVU10_MTOTZZA, IVU10_MTOTZZT)
resultados_IVU10_MTOT<-tabla(final)
resultados_IVU10_MTOT$nome<-Recode(resultados_IVU10_MTOT$nome,
                                       "'IVU10_MTOTZZC'='IVU10_MTOT&ZZC',
                                       'IVU10_MTOTZZS'='IVU10_MTOT&ZZS';
                                       'IVU10_MTOTZZA'='IVU10_MTOT&ZZA';
                                       'IVU10_MTOTZZT'='IVU10_MTOT&ZZT' " )

#####
#### IVU10_XTOT #####
#####

datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/IVU10_XTOT")
datos.new=ts(datos_web[, "dato"], start=c(1995,1), frequency = 12)

especificaciones<-spc(#usrdef.outliersEnabled = TRUE,
                      usrdef.outliersType = c(NA),
                      usrdef.outliersDate = c(NA),
                      #usrdef.varEnabled = TRUE,
                      #usrdef.var = regresseurs_J0,
                      #easter.type = "IncludeEaster",
                      #outlier.enabled = FALSE,
                      #automdl.enabled = FALSE,
                      #arima.mu= FALSE,

```

```

arima.p = 2, arima.d=1, arima.q = 0,
arima.bp = 0, arima.bd=1, arima.bq = 1)

ts_model <- tramoseats(datos.new, spec = especificaciones)

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
IVU10_XTOT<-series_datos[, "y"]

### Serie correxida de calendario:
IVU10_XTOTZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
deterministas[, "omhe"])

### Serie correxida de calendario e estacionalidade:
IVU10_XTOTZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
IVU10_XTOTZZA<-s_correx_S_C_AO_mensual(datos.new,ts_model)

### Serie de tendencia:
IVU10_XTOTZZT<-series_datos[, "t"]

final<-cbind(IVU10_XTOT, IVU10_XTOTZZC, IVU10_XTOTZZS, IVU10_XTOTZZA, IVU10_XTOTZZT)
resultados_IVU10_XTOT<-tabla(final)
resultados_IVU10_XTOT$nome<-Recode(resultados_IVU10_XTOT$nome,
                                       "'IVU10_XTOTZZC'='IVU10_XTOT&ZZC';
                                       'IVU10_XTOTZZS'='IVU10_XTOT&ZZS';
                                       'IVU10_XTOTZZA'='IVU10_XTOT&ZZA';
                                       'IVU10_XTOTZZT'='IVU10_XTOT&ZT'" )

#####
#### IVU10_MNE #####
#####

datos_web<-read.csv("http://www.ige.eu/igebdt/igeapi/csv/datosserc/IVU10_MNE")
datos.new<-ts(datos_web[, "dato"], start=c(1995,1), frequency = 12)

especificaciones<-spc(#usrdef.outliersEnabled = TRUE,
                      usrdef.outliersType = c("AO", "LS"),
                      usrdef.outliersDate = c("1999-11-01", "2000-09-01"),
                      #usrdef.varEnabled = TRUE,
                      #usrdef.var = regresseurs_JO,
                      #easter.type = "IncludeEaster",
                      #outlier.enabled = FALSE,
                      #automdl.enabled = FALSE,
                      #arima.mu= FALSE,
                      arima.p = 0, arima.d=1, arima.q = 1,
                      arima.bp = 0, arima.bd=1, arima.bq = 1)

ts_model <- tramoseats(datos.new, spec = especificaciones)

```

```

## Series correxidas ##

series_datos<-ts_model$final$series
deterministas<-ts_model$regarima$model$effects

### Serie de datos brutos:
IVU10_MNE<-series_datos[, "y"]

### Serie correxida de calendario:
IVU10_MNEZZC<-series_datos[, "y"]-(deterministas[, "tde"]+deterministas[, "ee"]+
deterministas[, "omhe"])

### Serie correxida de calendario e estacionalidade:
IVU10_MNEZZS<-series_datos[, "sa"]

### Serie correxida de calendario, estacionalidade e atípicos aditivos:
IVU10_MNEZZA<-s_correx_S_C_AO_mensual(datos.new,ts_model)

### Serie de tendencia:
IVU10_MNEZT<-series_datos[, "t"]

final<-cbind(IVU10_MNE, IVU10_MNEZZC, IVU10_MNEZZS, IVU10_MNEZZA, IVU10_MNEZT)
resultados_IVU10_MNE<-tabla(final)
resultados_IVU10_MNE$nome<-Recode(resultados_IVU10_MNE$nome,
                                     "'IVU10_MNE'='IVU10_MNE';
                                     'IVU10_MNEZZC'='IVU10_MNE&ZZC';
                                     'IVU10_MNEZZS'='IVU10_MNE&ZZS';
                                     'IVU10_MNEZZA'='IVU10_MNE&ZZA';
                                     'IVU10_MNEZT'='IVU10_MNE&ZT'" )

#####
resultados<-list(resultados_AMSS, resultados_PAROREG, resultados_EPA11_OCUT,
                  resultados_EPA11_PART, resultados_EPA11_INA, resultados_IPI15XE,
                  resultados_PRD_VEH, resultados_CPPGASOLT, resultados_CPPGO,
                  resultados_GEETO, resultados_ENRTO, resultados_ERTO,
                  resultados_VIINI, resultados_VITER, resultados_EDTVIV,
                  resultados_EDTSUP, resultados_SERV_CN_2015,
                  resultados_SERV_OCUP_2015, resultados_TAP, resultados_TMM,
                  resultados_EOHP_04, resultados_EOHVE_04, resultados_EOHGR_04,
                  resultados_EOTRP_04, resultados_EOTRVE_04, resultados_EOTRGR_04,
                  resultados_CPM_15_XERAL_09, resultados_IPRI15XE,
                  resultados_MATTUR, resultados_MATVCARGA, resultados_CXM,
                  resultados_CXX, resultados_IVU10_MTOT, resultados_IVU10_XTOT,
                  resultados_IVU10_MNE)

base_datos<-do.call(rbind,resultados)
base_datos$serie[base_datos$serie < 0] <- 0 # Poñer a cero os valores negativos

```

Bibliografía

- Box G. E. P., Cox D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, **26**, 2, 211-252.
- Box, G. E. P., Jenkins, G. M. (1976). *Time series analysis: Forecasting and control*. Holden-Day (2^a edición), San Francisco.
- Brockwell P. J., Davis R. A. (2002). *Introduction to Time Series and Forecasting*. Springer (2^a edición).
- Burnham K. P., Anderson D. R. (2004). Multimodel inference: understanding AIC and BIC in Model Selection. *Sociological Methods and Research*, **33**, 2, 261-304.
- Caporello G., Maravall A., Sanchez F. J. (2001). *Program TSW Reference Manual*. Banco de España.
- Dagum E. B. (1980). *The X-11-ARIMA seasonal adjustment method*. Statistics Canada.
- Departamento de Metodología y Desarrollo de la Producción Estadística (2019). *Manual de usuario de JDemetra+*. [Manual de software informático]. INE.
- Eurostat (2009). *ESS guidelines on seasonal adjustment : 2009 edition*. Eurostat.
- Eurostat, Directorate B (2020). Guidance on time series treatment in the context of the COVID-19 crisis.
- Friedman M. (1937). The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance. *Journal of the American Statistical Association*, **32**, 200, 675-701.
- Fujikoshi Y. (1993). Two-way ANOVA models with unbalanced data. *Discrete Mathematics*, **116**, 1, 315-334.
- Gómez V., Maravall A. (1996). *Programs Tramo and Seats. Instructions for the user (Beta version: September 1996)*. Banco de España.
- Grudkowska S. (2015). *JDemetra+ User Guide*. Narodowy Bank Polski, Department of Statistics, Polonia.
- Grudkowska S. (2016). *JDemetra+ Reference Manual Version 2.1*. Narodowy Bank Polski, Department of Statistics, Polonia.
- Instituto Galego de Estatística (2018). Notas explicativas. Banco de series de concuntura. Ciclotendencia e series corrixidas de estacionalidade e calendario. https://www.ige.gal/dba/metodoloxia_series_2018.pdf.
- Instituto Galego de Estatística (2022). *Banco de series do IGE*. <https://www.ige.eu/dba/Controlador?operacion=mostrarseries>. (Visitado: 2022-06-02)
- Instituto Nacional de Estadística (2019). *Estándar del INE para la corrección de efectos estacionales y efectos de calendario en las series coyunturales*. INE.
- Kruskal W. H., Wallis W. A. (1952). Use of Ranks in One-Criterion Variance Analysis. *Journal of the American Statistical Association*, **47**, 260, 583?621.

- Ladiray D., Quenneville B. (2000). *Methodologica. Revue des techniques, méthodes et instruments de recherche en sciences humaines. Número especial: Desestacionalizar con el método X-11*. Bruxelles: U.L.B. Laboratoire de méthodologie du traitement des données, Bélgica.
- Ljung G. M., Box G. E. P. (1978). On a Measure of a Lack of Fit in Time Series Models. *Biometrika*, **65**, 2, 297-303.
- Maravall A. (1987). Descomposición de series temporales. Especificación, estimación e inferencia., *Estadística Española*, **29**, 114, 11-69.
- Maravall A. (2012). *Update of Seasonality Tests and Automatic Model Identification in TRAMO-SEATS*. Banco de España.
- Neuhäuser M. (2011). *Wilcoxon?Mann?Whitney Test*. In: Lovric, M. (eds) *International Encyclopedia of Statistical Science*.. Springer, Berlin, Heidelberg.
- Novás M., Aparicio-Pérez F., López R., Saldaña S., Salgado D. e Sanguiaro-Sande L. (2021). Fixing the model for the seasonal component: A new revision policy. *Statistical Journal of the IAOS*, **37**, 4, 1229-1235.
- Palate J. (2022a). rjd3modelling: Interface to 'JDemetra+ 3.0' Seasonal Adjustment Software. <https://rdrr.io/github/palatej/rjd3modelling/>. (Visitado: 2022-06-02)
- Palate J. (2022b). rjd3toolkit: Interface to 'JDemetra+ 3.0' Seasonal Adjustment Software. <https://rdrr.io/github/palatej/rjd3toolkit/>. (Visitado: 2022-06-02)
- Peña D. (2005). *Análisis de series temporales*. Alianza Editorial.
- Planas C. (1997). The Analysis of Seasonality in Economic Statistics: A Survey of Recent Developments, *QUÈSTIÓ*, **22**, 1, 157-171.
- Quartier-la-Tente A., Michalek A., Palate J., Baeyens R. (2022). RJDemetra: Interface to 'JDemetra+' Seasonal Adjustment Software. [Manual de software informático] <https://cran.r-project.org/web/packages/RJDemetra/RJDemetra.pdf>.
- Ryan J. A., Ulrich J. M., Bennett R., Corwin J. (2020). xts: eXtensible Time Series. [Manual de software informático] <https://cran.r-project.org/web/packages/xts/xts.pdf>.
- Shiskin J., Young A. H., Musgrave J. C. (1967). *The X-11 variant of the census method II seasonal adjustment program*. Center for Statistical Research and Methodology. U.S. Census Bureau.
- Spinu V., Grolemund G.; Wickham H. (2021). lubridate: Make Dealing with Dates a Little Easier. [Manual de software informático] <https://cran.r-project.org/web/packages/lubridate/lubridate.pdf>.
- Tejedor F. J. (1999). *Ánalisis de varianza*. La Muralla S.A..
- Time Series Research Staff (2011). *X-12 ARIMA Reference Manual*. Center for Statistical Research and Methodology. U.S. Census Bureau.
- Time Series Research Staff (2015). *X-13 ARIMA-SEATS Reference Manual*. Center for Statistical Research and Methodology. U.S. Census Bureau.
- Villarreal F. G. (2005). *Elementos teóricos del ajuste estacional de series económicas utilizando X-12-ARIMA y TRAMO-SEATS*. Estudios Estadísticos 38, Naciones Unidas Comisión Económica para América Latina e o Caribe (CEPAL).