Department of Statistic Warsaw,
2016 r.

Sylwia Grudkowska – Narodowy Bank Polski, Department of Statistics  sylwia.grudkowska@nbp.pl,
(+48) 22 585 92 48

The views expressed herein are those of the authors and not necessarily those of the Narodowy Bank Polski

Print: NBP
Printshop
Published by:
Narodowy Bank Polski Education &
Publishing Department ul.
Świętokrzyska 11/21 00-919
Warszawa, Poland phone +48 22 653 23 35 www.nbp.pl

© Copyright Narodowy Bank Polski, 2016
This version of the *JDemetra+ User Manual* refers to JDemetra+ 2.1.0.

1. Introduction 7
   1.1. Historical background 7
   1.2. About JDemetra+ 10
      1.2.1. Installing and uninstalling JDemetra+ 12
      1.2.2. Running JDemetra+ 13
      1.2.3. Closing JDemetra+ 13
   1.3. About JDemetra+ Reference Manual 15
      1.3.1. Users to whom this document is intended 15
      1.3.2. How the document is organised 16
2. Main application windows 17
   2.1. Providers 17
   2.2. Workspace 20
   2.3. Results panel 22
3. Application menu 23
   3.1. File 23
   3.2. Statistical methods 24
      3.2.1. Anomaly Detection 24
      3.2.2. Modelling 29
      3.2.3. Seasonal Adjustment 30
   3.3. View 30
3.4. Tools 31
   3.4.1. Container 32
   3.4.2. Spectral analysis 38
   3.4.3. Aggregation 38
   3.4.4. Differencing 39
3.4.5. Spreadsheet profiler  40
3.4.6. Plugins  41
3.4.7. Options  42
3.5. Window  48
3.6. Help  50
3.7. RegArimaDoc  50
3.8. X13Doc  51
3.9. TramoDoc  52
3.10. TramoSeats doc  53
3.11. SAProcessingDoc  54
4. Modelling  58
4.1. Specifications  58
4.1.1. Tramo  59
4.1.2. RegARIMA  73
4.2. Documents  87
4.2.1. Input  91
4.2.2. Model - generalities  92
4.2.3. Forecasts  99
4.2.4. Regressors  101
4.2.5. Arima  102
4.2.6. Pre-adjustment series  103
4.2.7. Residuals  104
4.2.8. Likelihood  111
5. Seasonal adjustment  113
5.1. Specifications  113
5.1.1. TramoSeats  115
5.1.2. X13  119
5.2. Documents  125
5.2.1. TramoSeats  130
Acknowledgements:

I am deeply grateful to Agustin Maravall (Banco de España) and Dominique Ladiray (INSEE) for their invaluable guidance and contributions in the preparation of this document.

I am indebted to Duncan Elliott (ONS), Jennifer Davies (ONS), James Macey (ONS), Christiane Hofer (Bundesbank), Daniel Ollech (Bundesbank), Susanne Stollenmayer (Bundesbank), Anna Ciammola (ISTAT), Colm Bates (ECB) and for their insightful comments and suggestions.

I would like to thank all the members of the Seasonal Adjustment Centre of Competence for their useful comments and helpful suggestions on various drafts of this document.

Disclaimer:

The 'JDemetra+ Reference Manual' (2016) is provided by Eurostat. This material:

- Provides information to assist new users of JDemetra+ to familiarise themselves with the interface and functionalities of the application in a general nature and is not intended to favour one method over another that have been incorporated into the application;
- Is still in development;
- Sometimes links to further papers and documents for which Eurostat has no control and for which Eurostat assumes no responsibility;

Does not constitute professional or legal advice.

JDemetra+ is designed to support the ‘ESS Guidelines on Seasonal Adjustment’ (2015). While JDemetra+ incorporates the seasonal adjustment methods of the U.S. Bureau of Census (X-12-ARIMA and X-13ARIMA-SEATS) and the Banco de España (TRAMO/SEATS), the ‘ESS Guidelines on Seasonal Adjustment’ (2015) do not promote one method over another.

The paper presents the personal opinions of the author and does not necessarily reflect the official position of the institutions with whom the author cooperates. All errors are the author’s responsibility.

Contact:

Sylwia Grudkowska – Narodowy Bank Polski, Department of Statistics
sylwia.grudkowska@nbp.pl, (+48) 22 585 92 48
1. Introduction

1.1. Historical background

Seasonal adjustment (SA) is an important component of the official statistics business process. This technique is widely used for estimating and removing seasonal and calendar-related movements from time series resulting in data that present a clear picture of economic phenomena. For these reasons Eurostat\(^1\) takes part in various activities that aim to promote, develop and maintain a publicly available software solution for SA in line with established best practices.

Among many seasonal adjustment methods that produce reliable results for large datasets the most widely used and recommended ones are X-12-ARIMA/X-13ARIMA-SEATS\(^2\) developed at the U.S. Census Bureau and TRAMO/SEATS\(^3\) developed by Victor Gómez and Agustín Maravall, from the Banco de España. Both methods are divided into two main parts. The first part is called pre-adjustment and removes deterministic effects from the series by means of a regression model with ARIMA\(^4\) noise. The second part is the decomposition of the time series that aims to estimate and remove a seasonal component from the time series. TRAMO/SEATS and X-12-ARIMA/X13ARIMA-SEATS use a very similar approach in the first part to estimate the same model on the processing step, but they differ completely in the decomposition step. Therefore, comparing results

\(^1\) Eurostat is the statistical office of the European Union. Its task is to provide the European Union with statistics at European level that enable comparisons between countries and regions. More information is available at http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/.

\(^2\) X-12-ARIMA is a seasonal adjustment program developed and supported by the U.S. Census Bureau. It includes all the capabilities of the X-11 program (see DAGUM, E.B. (1980)) which estimates trend and seasonal component using moving averages. X-12-ARIMA offers useful enhancements including: extension of the time series with forecasts and backcasts from ARIMA models prior to seasonal adjustment, adjustment for effects estimated with user-defined regressors, additional seasonal and trend filter options, alternative seasonal-trend-irregular decomposition, additional diagnostics of the quality and stability of the adjustments, extensive time series modelling and model selection capabilities for linear regression models with ARIMA errors. For basic information on the X-12-ARIMA program see ’X-12-ARIMA Reference Manual’ (2011). More information on X-12ARIMA can be found at http://www.census.gov.

\(^3\) X-13ARIMA-SEATS is a seasonal adjustment program developed and supported by the U.S. Census Bureau that contains two seasonal adjustment modules: the enhanced X-11 seasonal adjustment procedure and ARIMA model based seasonal adjustment procedure from the SEATS seasonal adjustment program developed by GÓMEZ, V., and MARAVALL, A. (2013). For information on the X-13ARIMA-SEATS program see ’X-13ARIMA-SEATS Reference Manual’ (2015). More information on X13ARIMA-SEATS can be found at http://www.census.gov.

\(^4\) TRAMO/SEATS is a model-based seasonal adjustment method developed by Victor Gómez and Agustín Maravall (the Banco de España). It consists of two linked programs: TRAMO and SEATS. TRAMO (“Time Series Regression with ARIMA Noise, Missing Observations, and Outliers”) performs estimation, forecasting, and interpolation of regression models with missing observations and ARIMA errors, in the presence of possibly several types of outlier. SEATS (“Signal Extraction in ARIMA Time Series”) performs an ARIMA-based decomposition of an observed time series into unobserved components. Both programs are supported by the Banco de España. For basic information on the TRAMO/SEATS see CAPORELLO, G., and MARAVALL, A. (2004). More information on TRAMO/SEATS can be found at www.bde.es.

\(^5\) For description of the ARIMA model see 7.2.1.
from decomposition is often difficult. Furthermore, their diagnostics focus on different aspects and their outputs take completely different forms.

The TRAMO/SEATS method was originally implemented in 2001 in the program TSW (TramoSeats-Windows), which is a Windows extension of programs TRAMO and SEATS. Since then, a considerable amount of changes and additions have been added, that affect many important input parameters, as well as the output obtained. These changes resulted in program TSW+ launched in 2014. A LINUX version of TRAMO-SEATS is also available.

For X-13ARIMA-SEATS the U.S. Census Bureau provides the Windows interface called Win X-13. Distributions of X-13ARIMA-SEATS for Linux and Unix platforms are also available.

Both the above seasonal adjustment programs were originally written in FORTRAN, which is currently recognized as a declining language. The FORTRAN limitations - especially for the creation of reusable components and for the management of complex problems - make the maintenance of the relevant IT codes increasingly burdensome.

These original seasonal adjustment programs are commonly perceived by users as difficult to operate. Therefore, to improve access to these SA methods for non-specialists, Eurostat introduced new software called Demetra. It offered a user-friendly interface to the two SA algorithms: TRAMO/SEATS and X-12-ARIMA and facilitated the comparison of the output from those two algorithms. Even so, Demetra uses the FORTRAN libraries, which, together with an insufficient product development and handling of errors, is a factor that caused a rapid decline in software’s usage.

In 2009, the European Statistical System (ESS) launched its ‘Guidelines on Seasonal Adjustment’. As Demetra could not be adapted to the new requirements in the Guidelines, Eurostat, in cooperation with the National Bank of Belgium (NBB), started a project aiming to develop improved software called Demetra+. It was released in 2012. This tool provides a common approach for seasonal adjustment using the TRAMO/SEATS and X-12-ARIMA methods, which is more coherent with the Guidelines. It includes a unified graphical interface and input/output diagnostics for the two methods. Demetra+ source code is written in C++ and uses the two original FORTRAN modules.

---

7 Documentation on Win X-X13 can be found on: https://www.census.gov/srd/www/winx13/Winx13Doc.html.
9 Endorsed by the Statistical Programme Committee, the European Statistical System (ESS) ‘Guidelines on Seasonal Adjustment’ (2009) aim to harmonize European practices and to improve the comparability of infra-annual national statistics as well as enhance the overall quality of the European Union and the euro area aggregates. The ‘ESS Guidelines on Seasonal Adjustment’ (2009) and its revised version released in 2015 cover all the key steps of the seasonal and calendar adjustment process. They discuss both the theoretical aspects and practical implementation of seasonal adjustment issues.
as well as .NET libraries. Therefore Demetra+ software is non-extensible and cannot be used in IT environments other than Windows. For these reasons it seems that in long-term perspective it will not meet users’ expectations.

Therefore, Eurostat took an initiative to create new software that is based on Demetra+ experience but is platform independent and extensible. The resulting program is called JDemetra+. The NBB has been developing it since 2012. From the typical user perspective in comparison with Demetra+, numerous improvements have been implemented in JDemetra+, in terms of both layout and functionalities. But the most critical innovation is the re-writing of the original FORTRAN codes of X12-ARIMA/X-13ARIMA-SEATS and TRAMO/SEATS in JAVA, following a real object-oriented approach. These functionalities are discussed in the next section.
1.2. About JDemetra+

JDemetra+ is open source, platform independent, extensible software for seasonal adjustment (SA) and other related time series problems developed by the National Bank of Belgium. The tool includes the TRAMO/SEATS and X-12-ARIMA/X-13ARIMA-SEATS methods and enables the implementation of the ‘ESS Guidelines on Seasonal Adjustment’ (2015).

JDemetra+ offers up-to-date versions of leading seasonal adjustment algorithms rewritten in Java, which is a crucial factor that enables the long-term maintenance of the tool, integration of the libraries in the IT environments of many institutions and re-use of the modules and algorithms for other purposes. JDemetra+ is not only a user-friendly graphical interface, comparable to its predecessor, Demetra+, but also a set of open Java libraries that can be used to deal with time series related issues including the SA processing of large-scale datasets, the use of non-standard SA methods, the development of advanced research modules, temporal disaggregation, benchmarking and business cycle analysis. JDemetra+ is built around the concepts and the algorithms used in the two leading SA methods, i.e. TRAMO/SEATS and X-12-ARIMA/X-13ARIMA-SEATS. They have been re-engineered, following an object-oriented approach, which allows easier handling, extensions or modifications.

The most up-to-date version of JDemetra+ is JDemetra+ version 2.1.0, released in March 2016. In comparison with JDemetra+ version 2.0.0, it includes several enhancements: modification of an algorithm for allocation of AR roots in TRAMO/SEATS method and the calendarsigma and excludeforecast options in X-13-ARIMA-SEATS.

JDemetra+ version 2.1.0 is based on the following core engines:

- TramoSeats dlls, dated 10/2014;
- X12 dll (developed by the U.S. Census Bureau, based on X-12-ARIMA version 0.3, dated 12/2010).

One of the strategic choices for JDemetra+ is to provide common presentation/analysis tools for the seasonal adjustment methods included, so that the results from different methods can easily be compared. Obviously, JDemetra+ is highly influenced by the output of TRAMO/SEATS and of X12-ARIMA/X-13ARIMA-SEATS. Most analyses presented in JDemetra+ are available in the core engines. However, the results produced by JDemetra+ may slightly differ for several reasons (different statistical/algorithimc choices). In any case the global messages from seasonal adjustment are (nearly) always similar.

Among numerous important tools implemented in JDemetra+, the following functionalities should be highlighted:
RegARIMA modelling (using concepts developed in TRAMO and in X-12-ARIMA);
- Residuals analysis (mostly TRAMO-like);
- Seasonality tests (TRAMO and X-12-ARIMA-like);
- Spectral analysis (X-12-ARIMA definition);
- Sliding spans (X-12-ARIMA);
- Wiener-Kolmogorov analysis (for unobserved ARIMA components model, SEATS-like).

JDemetra+ is written using object-oriented programming (OOP) methodology. It allows developers to design software in a modular way, i.e. separate the functionality of an application into independent, interchangeable modules. Such units provide a specific group of functionalities and can be detached from the whole concept. The object-oriented approach is especially useful in the case of complex programs or when reusability matters.

Beside the statistical algorithms, JDemetra+ provides numerous peripheral services. The most important ones are the following:

- Dynamic access to various "time series providers": JDemetra+ provides modules to handle time series from different sources: Excel, databases (through JDBC), WEB services, files (TXT, TSW, USCB, XML, SDMX,...); the access is dynamic in the sense that time series are automatically refreshed, which consults the providers to download new information. The model allows asynchronous treatment.
- Common XML formatting: the seasonal adjustment processing can be saved in XML, which could also be used to generate WEB services around seasonal adjustment.

The graphical interface of JDemetra+ is based on the framework NetBeans\(^\text{11}\). Thanks to that technology external IT teams can create their own modules to enrich original software without modifying the core application. The main features that can be enriched are listed below:

Amongst the most important extension points, we have to mention:

- Time series providers: the users could add their own modules to download series from other databases;
- Diagnostics on seasonal adjustment;
- Output of SA processing.

As mentioned above, the API could be used to generate completely independent applications, but also to create, more easily, extensions to the current application.

\(^{11}\) See https://netbeans.org/.
One of the aims of JDemetra+ was to develop software which enables the comparison of the result from TRAMO/SEATS and X-12-ARIMA/X-13ARIMA-SEATS. For this reason, most of the analysis tools are common to both algorithms, e.g. the revision history and the sliding spans, even if they were originally developed in only one of them. On the other hand, all the features developed in the original programs have not always been implemented in JDemetra+; for instance, by contrast with TRAMO/SEATS, JDemetra+ does not separate the long-term trend from the cycle.

JDemetra+ runs on operating systems that support the Java VM (Virtual Machine) such as:

- Microsoft Windows Vista SP1/Windows 7/8;
- Ubuntu 9.10;
- Solaris OS version 11 Express (SPARC and x86/x64 Platform Edition);
- Macintosh OS X 10.6 Intel.

JDemetra+ runs on the Java SE Runtime Environment (JRE) version 7 update 21 or later. JRE can be downloaded to the user’s platform from one of the sites listed below:

- Windows, Linux and Solaris:
- Mac OS X: http://support.apple.com/downloads for Mac OS X 10.6

1.2.1. Installing and uninstalling JDemetra+

JDemetra+ is a stand-alone application packed in a zip package. It is accessible from Collaboration in research and methodology for official statistics webpage (www.cros-portal.eu).

Once downloaded to the PC it can be extracted to any folder on your system. The archive contains two versions of executable file (32-bit version and 64-bit version). The user should execute the version that matches the system version. The executable file is located in the appropriate nbdemetra/bin directory.

If the launching of JDemetra+ fails, you can try the following operations:

- Check if Java SE Runtime Environment (JRE) is properly installed by typing in the following command in a terminal: java --version
- Check the logs in your home directory:
  - %appdata%/.nbdemetra/dev/var/log/ for Windows;
  - ~/.nbdemetra/dev/var/log/ for Linux and Solaris;
  - ~/Library/Application Support/.nbdemetra/dev/var/log/ for Mac OS X.
In order to remove a previously installed JDemetra+ version, the user should delete an appropriate JDemetra+ folder.

1.2.2. Running JDemetra+

To open an application, navigate to the destination folder and double click on `nbdemetra.exe` or `nbdemetra64.exe` depending on the system version (`nbdemetra.exe` for the 32-bit system version and `nbdemetra64.exe` for the 64-bit system version).

![Figure 1.1: Running JDemetra+.](image)

1.2.3. Closing JDemetra+

To close the application, select `File → Exit` from the main menu (See Chapter 3).

![Figure 1.2: Closing JDemetra+.](image)
The other way is to click on the close box in the upper right-hand corner of the JDemetra+ window. If there is any unsaved work, JDemetra+ will display a warning and provide you with the opportunity to save it. The message box is shown below.

![Figure 1.3: The warning from leaving JDemetra+ without saving the workspace.](image)

---

Figure 1.3: The warning from leaving JDemetra+ without saving the workspace.
1.3. About JDemetra+ Reference Manual

The ‘JDemetra+ Reference Manual’ (2016) aims to introduce the user to the main features of JDemetra+ software and assist the user to take advantage of this powerful tool. The ‘JDemetra+ Reference Manual’ (2016) presents an overview of the capabilities of the software and includes detailed documentation on all of the specifications, with discussions of the available arguments and their default values. Also the results displayed by JDemetra+ are elaborated. Step-by-step descriptions of how to perform a typical analysis and useful tips that facilitate replication of the results with the user’s own data and working instructions are included in a separate document, namely the ‘JDemetra+ User Guide’ (2016).

JDemetra+ uses the notation "X12", "X13","Arima", "RegArima" and "TramoSeats" instead of "X-12ARIMA", "X-13ARIMA-SEATS", "ARIMA", "RegARIMA" and "TRAMO/SEATS" respectively. This notation is also used in the ‘JDemetra+ Reference Manual’ (2016) only when the references to the user interface are made.

1.3.1. Users to whom this document is intended

The ‘JDemetra+ Reference Manual’ (2016) can be used by beginners, i.e. those who have only basic knowledge about the idea of seasonality and its estimation in time series, as well as by users who have already acquired some experience from seasonal adjustment using relevant software and are able to interpret the outcomes, at least on a basic level.

This document does not describe the detailed working of the seasonal adjustment methods, nor the underlying mathematics. It assumes that the reader has acquired a background knowledge on the concept of seasonal adjustment and is familiar with the X-12-ARIMA, the X-13ARIMA-SEATS and the TRAMO/SEATS methods. A brief outline of the X-12-ARIMA, the X-13ARIMA-SEATS and the TRAMO/SEATS algorithms and concepts is included. For readers interested in studying the seasonal adjustment methods further, a bibliography is provided in Chapter 10.

It is assumed that the reader is familiar with concepts, such as stochastic and deterministic processes, time series, trend-cycle, seasonality, descriptive statistics, confidence level, mean square error, estimate, estimator, linear regression, stationarity, ARIMA process and so on. Readers with insufficient background to follow this document are encouraged to refer to an appropriate textbook, e.g. CHATFIELD (2004). Some background knowledge about seasonality in the time series can be gained from the e-learning courses on Seasonal Adjustment that are available at https://ec.europa.eu/eurostat/cros/search/custom-taxonomy/knowledge-repository-general-innovation-area/seasonal-adjustment.
1.3.2. How the document is organised

The ‘JDemetra+ Reference Manual’ (2016) is divided into seven parts, including an introduction and the Annex. The first two parts are designed to provide the user with necessary information about software and this document. The next chapters explore JDemetra+ capabilities and focus on functionalities that refer to the particular type of actions. They are further divided according to specific elements of the subject.

General information about JDemetra+ including an installation procedure is presented in Chapter 1. It also informs about the prerequisites and intended user’s profile.

Chapter 2 gives an overview of the software and explains the purpose of the main windows.

Chapter 3 discusses application menu and presents the available options.

An overview of the modelling features from the Workspace window is the focus of Chapter 4. It explains the options for the modelling specifications, explores the modelling facilities with TRAMO and RegARIMA models and gives insights into the output.

Chapter 5 discusses the options for the seasonal adjustment specifications and presents in detail the seasonal adjustment results from the pre-defined and user-defined specifications.

Chapter 6 elaborates on the calendars in JDemetra+ and user-defined regression variables. Both functionalities are designed to enhance the modelling capabilities of JDemetra+ and provide a better input for the decomposition stage.

The Annex describes selected aspects of the seasonal adjustment methods and technical issues including descriptions of the theoretical models used by X-12-ARIMA and TRAMO/SEATS. These descriptions of the seasonal adjustment related concepts are vital for a good understanding of the results produced by the software.

2. Main application windows

The default JDemetra+ window, which is displayed after launching the program, is clearly divided into several panels.
The key parts of the user interface are:

- The Providers window, which organises time series;
- The Workspace window, which stores results generated by the software as well as settings used to create them;
- A central empty zone for presenting the actual analyses further called the Results panel.

These areas are described in the next sections.

2.1. Providers

The Providers window presents the list of the data sources and organises the imported series within each data provider.
The allowed data sources include:

- JDBC;
- ODBC;
- SDMX;
- Spreadsheets;
- TSW;
- TXT;
- USCB; • XML.

All standard databases (Oracle, SQLServer, DB2, MySQL) are supported by JDemetra+ via JDBC which is a generic interface to many relational databases. Other providers can be added by users by creating plugins (see item 3.4). To import data, right-click on the appropriate provider from the Providers panel and specify the required parameters. For all providers the procedure follows the same logic. The requirements for each of these data sources as well as the importing procedure are discussed in the ‘JDemetra+ User Guide’ (2016), item 2.1.2.

The Providers window organises data in a tree structure reflecting the manner in which data are presented in the original source. The picture below presents how JDemetra+ visualises the imported spreadsheet file. If the user expands all the pluses under the spreadsheet all the series within each sheet that has been loaded are visible. Here two time series are visible: Japan (under the Asia branch) and United States (under the North America branch) while the Europe branch is still folded. The names of the time series have been taken from the column headings of the spreadsheet while the names of the branches come from sheets’ names.
2.3: Dataset structure.

Series uploaded to the Providers window can be displayed (see "JDemetra+ User Guide" (2016), item 2.1.4), modified and tested (item 3.4) and used in estimation routines (Chapters 4 and 5). The data sources can be restored after re-starting the application so that there is no need to fetch them again. This functionality can be set in the Behaviour tab available at the Option item from the Tools menu.

2.2. Workspace

Workspace is a JDemetra+ functionality that stores the work performed by the user in a coherent and structured way. By default, each workspace contains pre-defined modelling and seasonal adjustment specifications and basic calendar. A specification is a set modelling and/or seasonal adjustment parameters.

Within the workspace the following items can be saved:

- User-defined modelling and seasonal adjustment specifications;
- Documents, that contain results from time series modelling and seasonal adjustment;
- User-defined calendars;
- User-defined regression variables.

Together with the results from modelling and seasonal adjustment, also the original data, paths to the input files and parameters of processes are saved. Therefore, these results can be re-opened, updated, investigated and modified in next sessions with JDemetra+.

The workspace saved by JDemetra+ includes:
- Main folder containing several folders that correspond to the different types of items created by the user and;
- The xml file that enables the user to import the workspace to the application and to display its content.

An example is shown in Figure 2.4.

The workspace can be shared with other users, which eases the burden of work with defining specifications as well modelling and seasonal adjustment processes.

The content of the workspace is presented in the **Workspace** window. It is divided into three sections: **Modelling**, **Seasonal adjustment** and **Utilities**. The window organises all default and user-defined specifications (item 4.1 for the modelling specifications and 5.1 for the seasonal adjustment specifications), documents, which contain the results of the respective processes (see the item 4.2 for the modelling documents and 5.2 for the seasonal adjustment documents), calendars (item 6.1) and user defined variables (item 6.2).
2.3. Results panel

The empty zone on the right is the place where JDemetra+ displays various windows. More than one window can be displayed at the same time. Windows can overlap each other with the foremost window being in focus or active. The active window has a darkened title bar. The windows in the results panel can be arranged in many different ways, depending on the user’s needs (see item 3.5). The example below shows one of the possible views of this panel. The results of the user’s analysis are displayed in respective windows. The picture below shows two panels – a window containing seasonal adjustment results (upper panel) and the other one containing autoregressive spectrum (lower panel).

Figure 2.6: The Results panel filled with two windows.

3. Application menu

The majority of functionalities are available from the menu of the application, which is situated at the very top of the main window. If the user moves the cursor to an entry in the main menu and clicks on the left mouse button, a drop-down menu will appear. Clicking on an entry in the dropdown menu selects the highlighted item.
3.1. File

The File menu is intended to create new workspaces, open existing workspaces and datasets and saving new/updated workspaces.

![File menu](image)

Figure 3.2: The content of the File menu.

It offers the following functions:

- **New Workspace** – creates a new workspace and displays it in the Workspace window with default name (Workspace_#number);
- **Open Workspace** – opens the dialog window which enables the user to select and open an existing workspace;
- **Open Recent Workspace** – presents the list of workspaces recently created by the user and enables the user to open one of them;
- **Save Workspace** – saves the project file named by the system under the default name (Workspace_#number) and in the default location. The workspace can be re-opened at a later point in time;
- **Save Workspace As...** – saves the project file named by the user in the chosen location. The workspace can be re-opened at a later point in time;
- **Open Recent** – presents the list of datasets recently used and enables the user to open one of them;
- **Exit** – closes an application.

### 3.2. Statistical methods

The Statistical methods menu includes functionalities for modelling, analysis and seasonal adjustment of a time series. They are divided into three groups:

- **Anomaly Detection** – allows for purely automatic identification of regression effects;
- **Modelling** – enables time series modelling using the TRAMO and RegARIMA models;
- **Seasonal adjustment** – intended for seasonal adjustment of the time series with the TRAMO/SEATS and X-13ARIMA-SEATS methods.

![Statistical methods menu](image)

#### Figure 3.3: The Statistical methods menu.

### 3.2.1. Anomaly Detection

The primary goal of functionalities available in the Anomaly Detection section is an identification of atypical values called outliers. According to the ‘ESS Guidelines on Seasonal Adjustment’ (2015), seasonal adjustment methods are likely to be severely affected by the presence of such values; therefore they should be detected and replaced simultaneously or before estimating the seasonal and calendar components in order to avoid a distorted or biased estimation of them.

The use of the RegARIMA models is recommended by the ‘ESS Guidelines on Seasonal Adjustment’ (2015) to estimate and remove outliers before estimating the seasonal effect. As the presence of outliers could strongly affect the quality of the decomposition, the various types of outlier (i.e. additive outliers, transitory changes, level shifts, etc.) should be detected and corrected for.

The quality control should be performed each time new or revised data become available. The manual inspection of the data is problematic, especially in the case of large databases. Moreover, it usually relies on some simple measures which do not consider the full information contained in
the series, but just a few values. Therefore the results can be strongly affected by seasonality, noise, or special events.

JDemetra+ includes two tools dedicated for automatic identification of outliers: *Check Last* and *Outliers Detection*. Both are based on the TERROR program, which is an application of TRAMO, executed in an automatic manner (with several possible options to set) to the problem of quality control in time series for automatic detection of errors in time series. Therefore, both tools use the TRAMO specifications.

![Statistical methods menu](image)

Figure 3.4: The Anomaly detection menu.

3.2.1.1. Check Last

The *Check Last* tool automatically detects the TRAMO model, forecasts up to three last observations and marks the observations which are too different from the forecasted value. The *Check Last Batch* window is divided into three panels. The panel on the left presents the list of analysed series. The results are displayed in the panels on the right.

To launch the analysis, drag and drop the series from the *Providers* window into the left hand side panel of the *Check Last Batch* window and click the *Start* button (denoted with the green arrow) from the menu in the top part of the window. The analysis will be performed using the TRAMO specification selected from the list. By default, TR4 is used.

![Check Last initial window](image)

Figure 3.5: The Check Last initial window with series to be processed and the list of available specifications expanded.

JDemetra+ removes the last observations from the series and calculates a one-period-ahead out-of-sample forecast of the series. The forecasted values are then compared with the actual values. The user may decide how many of the last observations will be considered (one, two, or three) in
this procedure (click on the 123 button and specify the number). The number of columns visible in the panel on the left is adjusted accordingly to the user choice.

Figure 3.6: The options for number of observations to be examined.

The default settings can be changed in the Properties dialog box (the number of last observations that will be compared to the forecasted values, specification used for modelling and the threshold values used to decide if observation is abnormal). To open it, click on the button marked with the working tools.

Figure 3.7: The properties for the Check Last functionality.
Once the process is executed, click on the series on the list to display the results. For each series the program automatically identifies an ARIMA model, detects several types of outlier, interpolates for missing values and estimates calendar effects, if appropriate. Study the detailed results section using the vertical scrollbar.

![Figure 3.8: The results of the outlier's detection process.](image)

The last observations (one, two or three, depending of the user’s choice) are compared with the forecasted values. If, for a given observation, the forecast error divided by standard deviation of residuals is greater than the first threshold value and lower than the second threshold value then this observation is classified as containing a "possible error" and marked in orange. If this value is greater than the second threshold value than the new observation is classified as containing a "likely error" and marked in red. Otherwise, the observation is accepted as without error.\(^\text{12}\)

---

JDemetra+ enables the user to save the results of this analysis in the compact form of the report. To generate it, click on the Generate Report button and specify the sorting options.

To save the report click OK and select the destination folder.
3.2.1.2. Outlier detection

The Outlier Detection tool allows for identification of an ARIMA model, including detection of outliers, interpolation of missing values and estimation of calendar effects. The assessment of the last observations’ quality is not performed. Instead, the outliers are highlighted in green (transitory changes), red (level shifts), blue (additive outliers) and violet (seasonal outliers). By default the seasonal outliers are not considered. The user may change this setting in the Properties menu.

Step-by-step demonstration of the Outlier Detection tool capabilities can be found in the ‘JDemetra+ User Guide’ (2016), item 3.3.1.

3.2.2. Modelling

The aim of the Modelling section is to provide tools for time series modelling and forecasting without performing the estimation of the components and decomposition. The estimation results can be useful for time series analysis and prediction of the short-term development.

The Modelling section includes all capabilities from the TRAMO and RegARIMA models. It is flexible in specifying model parameters. The results can be saved and refreshed with updated series.

Instructions for use this functionality is given in the ‘JDemetra+ User Guide’ (2016), item 3.3.2.
3.2.3. Seasonal Adjustment

The Seasonal Adjustment section provides tools to perform seasonal adjustment for a single time series as well as for multiple time series using the TRAMO/SEATS or X-13ARIMA-SEATS methods. It also offers several seasonality tests that can be used to scrutinize the presence and the nature of seasonal movements in time series independently from seasonal adjustment. Finally, the DirectIndirect Seasonal Adjustment tool enables for comparison the results from direct and indirect seasonal adjustment performed for the aggregated series.

The guidance for using these functionalities is given in the ‘JDemetra+ User Guide’ (2016) (basic scenario that allows for generating seasonal adjustment in an automatic way is presented in 3.1, different types of user interventions are discussed in 3.2, the Seasonality Tests tool is explained in 3.4.1 while the capabilities of Direct-Indirect Seasonal Adjustment are shown in 3.2.1.8).

3.3. View

The View menu contains functionalities allowing for customising the view of an application to the user needs. It offers the following items:

- **Split** – the function is not operational in the current version of the software.
- **Toolbars** – displays selected toolbars under main menu. The File toolbar contains the Save all icon. The Performance toolbar includes two icons: one to show the performance of the application, the other to stop application profiling and take a snapshot. The Other toolbar determines the default behaviour of the program when the user double clicks on the data. It may be useful to plot the data, visualise it on a grid, or to perform any pre-specified action, e.g. execute a seasonal adjustment procedure.
- **Show Only Editor** – displays only the Results panel and hides other windows (e.g. Workspace and Providers).
- **Full Screen** – displays the current JDemetra+ view in the full screen.
3.4. Tools

The Tools menu includes, among other functionalities, tools helpful for graphical analysis of the time series. The ‘X-13ARIMA-SEATS Reference Manual’ (2015) strongly recommends studying a high resolution plot of the time series as it is helpful to get information about e.g. seasonal patterns, potential outliers and stochastic non-stationarity. Also the ‘ESS Guidelines on Seasonal Adjustment’ (2015) recommend carrying out a graphical analysis on both unadjusted data and the initial run of the seasonal adjustment software. In particular, the graphical analysis should consider:

- The length of the series and model span;
- The presence of zeros or outliers or problems in the data;
- The structure of the series: presence of a long term and cyclical movements, of a seasonal component, volatility etc.;
- The presence of possible breaks in the seasonal behaviour;
- The decomposition scheme (additive, multiplicative).

The ‘ESS Guidelines on Seasonal Adjustment’ (2015), recommend that this exercise should be performed and documented for the most important series to be adjusted at least once a year.

The following functionalities are available from the Tools menu:

- **Container** – includes several tools for displaying data in time domain;
- **Spectral analysis** – contains tools for analysis of time series in a frequency domain;
- **Aggregation** – enables the user to investigate the graph of the sum of multiple time series;
- **Differencing** – allows for inspection the first regular differences of the time series;
- **Spreadsheet profiler** – offers the Excel-type view of the XLS file imported to JDemetra+.
- **Templates** – an option not used in current version of JDemetra+.

- **Plugins** – allows the installation and activation of plugins, which extend JDemetra+ abilities.

- **Options** – presents default interface settings and allows for their modification.

![Figure 3.15: The Tools menu.](image)

### 3.4.1. Container

*Container* includes basic tools to display the data. The following items are available: *Chart, Grid, Growth Chart* and *List*.

![Figure 3.16: The Container tools.](image)

Several containers can be opened at the same time. Each of them may include multiple time series.

*Chart* plots the time series as a graph. This function opens an empty window. To display a given series drag and drop the series from the *Providers* window into it. More than one series can be displayed on one graph.
Figure 3.17: Launching the Chart functionality.

The series to display can be also dragged from other windows (e.g. the Variables window) or directly from windows that display the results of the estimation procedure.

Figure 3.18: Displaying the seasonally adjusted series on a separate chart.

To adjust the view of the chart and save it in a given location use the local menu, which is displayed after right-clicking on the chart. The explanation for the functions available for the local menu is given below Figure 3.20.
To display the time series value at a given date, hover over it with the cursor. Once the time series is marked by clicking on it with the right mouse button, additional options are available.

- **Open** – opens selected time series in the new window that contain Chart and Grid panels.
- **Open with** – opens the time series in a separate window according to the user choice (chart & grid or only chart). The All ts views option is not currently available.
- **Rename** – enables the user to change the time series name.
- **Freeze** – disables modifications of the chart.
- **Copy** – copies series and allows it to be pasted into another application e.g. into Excel.
- **Paste** – pastes the time series previously marked.
- **Remove** – removes a time series from the chart.
- **Split into yearly components** – opens a window that presents the analysed series data split by year. This chart is useful to investigate the differences in time series values caused by the seasonal factors as it gives some idea about the existence and size of deterministic and stochastic seasonality in data.
- **Select all** – selects all time series presented in the graph.
- **Clear** – removes all time series from the chart.
- **Show title** – option is not currently available
- **Show legend** – displays titles of all time series presented on the graph.
- **Edit format** – enables the user to change data format.
- **Color scheme** – allows the colour scheme used in the graph to be changed.
- **Lines thickness** – allows the user to choose between thin and thick lines to be used for a graph.
- **Show all** – this option is not currently available.
- **Export image to** – allows the graph to be sent to the printer and saved as clipboard or as a file in the PNG format.
- **Configure** – enables the user to customize chart and series display.

Some of these options are investigated further in the ‘*JDemetra+ User Guide*’ (2016), 2.1.4.

*Grid* displays the selected time series as a table. This function opens an empty window. To display a given series drag and drop the series from the *Providers* window into it. More than one series can be displayed in one table. The chart is automatically rescaled after adding a new series.

To display available options, left click on the series header to select the series and then right click on any time series data.
Options that are characteristic for Grid are:

- **Transpose** – changes the orientation of the table from horizontal to vertical.
- **Reverse chronology** – displays the series from the last to the first observation.
- **Single time series** – removes from the table all time series apart from selected one.
- **Edit format** – allows the format used for displaying dates and values to be changed.
- **Use color scheme** – allows the series to be displayed in colour.
- **Color scheme** – allows for a choice of color scheme from a pre-specified list.
- **Show bars** – presents a value in table as horizontal bars.
- **Zoom** – option for modifying the chart size.

The explanation of other options can be found below Figure 3.20.

**Growth chart** presents year-over-year or period-over-period growth rates for selected time series. This function opens an empty window. To display a transformation for a given series, drag and drop the series from the **Providers** window into it. More than one series can be displayed in one table. The growth chart is automatically rescaled after adding a new series.
Left click displays the local menu with available options. These that are characteristic for Growth chart are:

- **Kind** – displays m/m (or q/q) and y/y growth rates for all time series in the chart (previous period and previous year options respectively). By default, the period-over-period growth rates are shown.
- **Edit last year** – for clarity and readability purposes, by default only five last years of observations are shown. However, this parameter can be adjusted to the user needs.

The explanation of other options can be found below Figure 3.20.

*List* provides basic information about chosen time series, like start and end date, number of observations and sketch of data graph. This function opens an empty window. To display information, drag and drop the series from the *Providers* window into the *List* window. Right click displays the local menu with available options. Apart from standard options, the local menu for *List* enables marking the series that match the selected frequency (yearly, half-yearly, quarterly, monthly) using the *Select by frequency* option. The explanation of other options can be found below Figure 3.20.
3.4.2. Spectral analysis

The Spectral analysis section provides three spectral graphs for in-depth analysis of time series in the frequency domain. These graphs are Auto-regressive Spectrum, Periodogram and Tukey Spectrum. Brief information concerning spectral analysis is given in 7.3. A detailed presentation of the abovementioned tools can be found in the ‘JDemetra+ User Guide’ (2016), item 3.4.2.

3.4.3. Aggregation

Aggregation calculates the sum of the selected series and provides basic information about the selected time series, including the start and end date, number of observations and a sketch of the data graph, in the same way as the List functionality. Aggregation opens an empty window. To sum the selected series, drag and drop them from the Providers window into the Aggregation window. Right click displays the local menu with available options. The content of the local menu depends
on the panel chosen (panel on the left that contains the list of the series or panel on the right that presents the graph of an aggregate). The local menu for the list of series offers the option Select by frequency that marks all the series on the list that are yearly, half-yearly, quarterly or monthly (depending on the user’s choice). The explanation of other options can be found below Figure 3.20.

3.4.4. Differencing

The Differencing window displays the first regular differences for selected time series together with the corresponding periodogram and the PACF function. By default, the window presents the results for non-seasonally and seasonally differenced series \((d = 1, D = 1)\). These settings can be changed through the Properties window (Tools → Properties).

Figure 3.26: The Aggregation tool.

Figure 3.27: The properties of the Differencing tool.
The typical results are shown below. The bottom left graph presents the partial autocorrelation coefficients (vertical bars) and the confidence intervals. The right-click local menu offers several functionalities for differenced series. The explanation of available options can be found below Figure 3.20.

For the Partial autocorrelation and the Periodogram panels the right-button menu offers a copy series option that allows data to be exported to another application and a graph to be printed and saved as a clipboard or as a PNG file. Information about partial autocorrelation function is given in 7.9. The description of periodogram is available in the ‘JDemetra+ User Guide’ (2016), item 3.4.2

3.4.5. Spreadsheet profiler

The Spreadsheet profiler offers the Excel-type view of the XLS file imported to JDemetra+. To use this functionality drag the file name from the Providers window and drop it the empty Spreadsheet profiler window.
3.4.6. Plugins

JDemetra+ is an application that supports plug-ins. A plugin is a software component that adds a specific feature to an existing software application and is independent from the software version. It allows an application to be enhanced without changing the original code. Plugins can be shared between the users and installed individually. In this way new functionalities can be easily distributed among seasonal adjustment community members.

The Plugins window includes five panels: Updates, Available plugins, Downloaded, Installed and Settings, some of them however are not operational in the current version of the software.

The aim of the Updates panel is to allow the user for the manual check if some updates of the already installed plugins are available. This functionality, however, is currently not operational for the JDemetra+ plugins.

The Available plugins panel allows the downloading of all plugins that are related to JDemetra+. This functionality, however, is currently not operational for the JDemetra+ plugins.

The Downloaded panel is designed for installation of the new plugins from the local machine. This process in explained in detail in 7.10.

The Settings panel is designated for adding the update centres, which are the places to hold plugins. For each centre the user can specify proxy settings and a time interval for the automatic checking for updates. At the moment this functionality is not operational for the JDemetra+ plugins.
3.4.7. Options

The Options window includes five main panels: Demetra, General, Keymap, Appearance and Miscellaneous. They are visible in the very top of the Options window.

By default, the Demetra tab is shown. It is divided into seven areas: Behaviour, Demetra UI, Statistics, Data transfer, Demetra Paths, ProcDocumentItems, and Interchange.

Behaviour defines the default reaction of JDemetra+ to some of the actions performed by the user.

- **Providers** – an option to show only the data providers that are available.
- **Persistence** – an option to restore the data sources after re-starting the application so that there is no need to fetch them again. It is also possible to restore all the content of chart and grid tools.
- **Threading** – defines how resources are allocated to the computation (Batch Pool Size controls the number of cores used in parallel computation and Batch Priority defines the priority of computation over other processes). Changing these values might improve computation speed but also reduce user interface responsiveness.
- **Time Series** – determines the default behaviour of the program when the user double clicks on the data. It may be useful to plot the data, visualise it on a grid, or to perform any prespecified action, e.g. execute a seasonal adjustment procedure.
The Demetra UI tab allows for setting:

- Default colour scheme for the graphs.
- Data format (uses MS Excel conventions). For example, \#,\#,\#,\# implies the numbers in the tables and the y-axis of the graphs will be rounded up to four decimals after the point.
- Default number of last years of the time series displayed in charts representing growth rates.
- Visibility of icons in the context menus.
The Statistics tab includes options to control:

- The number of years used for spectral analysis and for model stability;
- The default pre-defined specification for seasonal adjustment;
- The settings for quality measures and tests used in a diagnostic procedure; and
- The type of the analysis of revision history.
To modify the settings for a particular measure, double click on a selected row, introduce changes in a pop-up window and click OK. The description of the parameters for each quality measure and test used in a diagnostic procedure can be found in 7.12.
The Data transfer tab contains the multiple technologies that defined the behaviour of the drag and drop and copy-paste actions. To change the default settings, double click on the selected item. Once the modifications are introduced, confirm them with the OK button.

Figure 3.35: The content of the Data Transfer tab.
*Demetra paths* allows the user to specify the relative location of the folders where the data can be found. This way, the application can use the data from different computers. Otherwise, the user would need to have access to the exact path where the data is located. To add a relative location, select the data provider, click “+” button and specify the location.

![Image of Demetra Paths tab](image)

**Figure 3.36:** The content of the *Demetra Paths* tab.

*ProcDocumentItems* includes a list all reports available for processed documents like seasonal adjustment. The Interchange tab is not operational in the current version of the software.

The next section, *General*, allows for customise a proxy settings. A proxy is an intermediate server that allows an application to access the Internet. It is typically used inside a corporate network where Internet access is restricted. In JDemetra+, the proxy is used to get time series from remote servers like .Stat.
Keymap provides a list of default key shortcuts to access some of the functionalities and it allows the user to define others.

The Appearance and Miscellaneous tabs are tabs automatically provided by the Netbeans platform. They are not used by JDemetra+.

3.5. Window

The Window menu offers several functions that facilitate the analysis of data and enables the user to adjust the interface view to the user’s needs.
Figure 3.39: The **Window** menu.

- **Preview Time Series** – opens a window that plots any of the series the user selects from **Providers**.
- **Debug** – opens a **Preview Time Series** window that enables for fast display the graphs for time series from a large dataset. To display the graph click on the series in the **Providers** window.
- **Providers** – opens (if closed) and activates the **Providers** window.
- **Variables** – opens (if closed) and activates the **Variable** window.
- **Workspace** – opens (if closed) and activates the **Workspace** window.
- **Output** – generic window to display outputs in the form of text; useful with certain plugins (e.g. tutorial descriptive statistics).
- **Editor** – activates the editor panel (and update the main menu consequently).
- **Configure Window** – enables the user to change the way of display the window (maximize, float, float group, minimize, minimize group). The options are active when some window is activated.
- **Properties** – opens the **Properties** window and displays the properties of the marked item (e.g. time series, data source);
- **Reset Windows** – restores the default JDemetra+ view.
- **Close Window** – closes all windows that are open.
- **Close All Documents** – closes all documents that are open.
• **Close Other Documents** – closes all documents that are open except for the one which is active (is the last activated one).
• **Document Groups** – enables the user to create and manage the document groups.
• **Documents** – lists all documents that are active.

### 3.6. Help

The *Help* menu allows access to the documentation attached to software (*Help Contents*). For JDemetra+ 2.1.0 no documentation is provided. *Check for Updates* activates the *Plugin installer* that downloads, verifies and installs the selected plugins, if available. *About* includes basic information about JDemetra+ and Java versions and directories used by an application.

![Help Menu](image)

Figure 3.40: The *Help* menu.

### 3.7. RegArimaDoc

*RegArimaDoc* is a main menu item that includes functionalities designated for *RegArima* documents. This item is displayed in the main menu of the application when a newly created or existing *RegArima* document is active.

![RegArimaDoc Menu](image)

Figure 3.41: The *RegArimaDoc* menu for a newly created *RegArima* document

*RegArimaDoc-*#number* is a default name that corresponds to the *RegArima* document’s name. It is updated automatically, once the relevant *RegArima* document is renamed.
The options for *RegArimaDoc* include:

- **Paste** – pastes the time series to the *RegArimaDoc* window and performs seasonal adjustment for this time series using settings selected for the current process. The series need to be previously copied, e.g. from the *Providers* window. Otherwise the message informing that the action cannot be done is displayed.

- **Refresh data** – updates the input data and executes the seasonal adjustment process. This option is only available for *RegArimaDoc* documents saved during a previous session with JDemetra+. To use this option, first create *RegArimaDoc* document and drag and drop the time series into it. Then save the workspace and close JDemetra+. Next, update the time series (add/change the observations, but do not change neither the localisation of the file nor the file name) and open JDemetra+ once again. Open the previously saved workspace and double click on the respective *RegArima* document in the *Workspace* window. Finally, chose the **Refresh data** option from the *RegArimaDoc* menu. JDemetra+ re-estimates the complete seasonal adjustment model automatically, so the results are updated immediately. The user can modify the specification and validate the newly introduced changes using the *Specifications* functionality (see 4.1).

- **Lock** – prevents the series from the *RegArimaDoc* from being changed when the user double clicks on a new one in the *Providers* window. However, with this option the user can still change specification’s settings.

### 3.8. X13Doc

*X13Doc* is a main menu item that includes functionalities designated for *X13* documents. This item is displayed in the main menu of the application when a newly created or existing *X13* document is active.
X13Doc-#number is a default name that corresponds to the X13 document’s name. It is updated automatically, once the relevant X13 document is renamed.

The options for the X13Doc include:

- **Paste** – pastes the time series to the X13Doc window and performs seasonal adjustment for this time series using settings selected for the current process. The series need to be previously copied, e.g. from the Providers window. Otherwise the message informing that the action cannot be done is displayed.

- **Refresh data** – updates the input data and executes the seasonal adjustment process. This option is only available for X13 documents saved during a previous session with JDemetra+. To use this option, first create X13 document and drag and drop the data into it. Then save the workspace and close JDemetra+. Next, update the time series (add/change the observations, but do not change either the localisation of the file or the file name) and open JDemetra+ once again. Open the previously saved workspace and double click on the respective X13 document in the Workspace window. Finally, chose the Refresh data option from the X13Doc menu. JDemetra+ re-estimates the complete seasonal adjustment model automatically, so the results are updated immediately. The user can modify the specification and validate the newly introduced changes using the Specification functionality (see 4.1 and 5.1).

- **Lock** – prevents the series from the X13Doc from being changed when the user double clicks on a new one in the Providers window. However, with this option the user can still change the specification’s settings.
3.9. TramoDoc

*TramoDoc* is a main menu item that includes functionalities designated for *Tramo* documents. This item is displayed in the main menu of the application when a newly created or existing *Tramo* document is active.

![TramoDoc menu](image)

*TramoDoc*-#number is a default name that corresponds to the *Tramo* document’s name. It is updated automatically, once the relevant *Tramo* document is renamed.

The options for the *TramoDoc* include:

- **Paste** – pastes the time series to the *TramoDoc* window and performs seasonal adjustment for this time series using settings selected for the current process. The series need to be previously copied, e.g. from the *Providers* window. Otherwise the message informing that the action cannot be done is displayed.

- **Refresh data** – updates the input data and executes the seasonal adjustment process. This option is only available for *Tramo* documents saved during previous session with JDemetra+. To use this option, first create *Tramo* document and drag and drop the data into it. Then save the workspace and close JDemetra+. Next, update the time series (add/change the observations, but do change neither the localisation of the file nor the file name) and open JDemetra+ once again. Open the previously saved workspace and double click on the respective *Tramo* document in the *Workspace* window. Finally, chose the **Refresh data** option from the *TramoDoc* menu. JDemetra+ re-estimates the complete seasonal adjustment model automatically, so the results are updated immediately. The user can modify the specification and validate the newly introduced changes using the *Specification* functionality (see 4.1).
• **Lock** – prevents the series from the *TramoDoc* from being changed when the user double clicks on a new one in the *Providers* window. However, with this option the user can still change specification’s settings.

### 3.10. TramoSeats doc

*TramoSeatsDoc* is a default name for the main menu item that includes functionalities designated for *TramoSeats* documents. This item is displayed in the main menu of the application when a newly created or existing *TramoSeats* document is active.

![Figure 3.44: The *TramoSeatsDoc* menu for a newly created X13 document.](image)

*TramoSeatsDoc* is a default name that corresponds to the *TramoSeats* document’s name. It is updated automatically, once the relevant *TramoSeats* document is renamed.

The options for the *TramoSeatsDoc* include:

- **Paste** – pastes the time series to the *TramoSeatsDoc* window and performs seasonal adjustment for this time series using settings selected for the current process. The series need to be previously copied, e.g. from the *Providers* window. Otherwise the message informing that the action cannot be done is displayed.

- **Refresh data** – updates the input data and executes the seasonal adjustment process. This option is only available for *TramoSeats* documents saved during a previous session with JDemetra+. To use this option, first create *TramoSeats* document and drag and drop the data into it. Then save the workspace and close JDemetra+. Next, update the time series (add/change the observations, but do change neither the localisation of the file nor the file name) and open JDemetra+ once again. Open the previously saved workspace and double click on the respective *TramoSeats* document in the *Workspace* window. Finally, chose the **Refresh data** option from the *TramoSeatsDoc* menu. JDemetra+ re-estimates the complete
seasonal adjustment model automatically, so the results are updated immediately. The user can modify the specification and validate the newly introduced changes using the Specification functionality (see 4.1 and 5.1).

- **Lock** – prevents the series from the TramoSeatsDoc from being changed when the user double clicks on a new one in the Providers window. However, with this option the user can still change specification’s settings.

### 3.11. SAProcessingDoc

SAProcessingDoc-\#number is a main menu item that includes functionalities designated for multidocuments. This item is displayed in the main menu of the application when a newly created or existing multi-document is active. Note that some options are inactive when none of the time series in the SAProcessing window is marked.

![Figure 3.45: The SAProcessing menu for a newly created X13 document.](image)

The SAProcessingDoc-\#number menu facilitates the management of the multi-document.
The available options include:

- **Default specification** – opens the window which contains the list of available specifications (both pre-defined and user-defined).
- **Start** – runs the seasonal adjustment of the times series included in the document.
- **Refresh** – refreshes a document with new data using one of the revisions policies. This option is only available for multi-documents saved during a previous session with JDemetra+. The description of refresh policies and step-by-step demonstration of how to use them can be found in the ‘JDemetra+ User Guide’ (2016), item 3.2.2.2.

- **Accept** – for a marked series the option allows the user to accept the results which JDemetra+ marks as not satisfactory. However, the option changes the entry in the *Quality* column in the *SAProcessing* window into **Accepted**, regardless of the initial value of this field. The **Accept** function can be used to distinguish the series for which the results have been examined by the user from those not checked yet. The option can be also activated from the local menu. To use it, first the series in the *SAProcessing* window need to be marked as it is shown in the picture below.
Figure 3.47: The impact of the Accept option on the view of the results in the multi-document.

- **Edit** – allows for modification of the content of multi-document. New time series can be added to the multi-document directly from the external source, e.g. Excel (Paste). Before choosing this option the user should copy the time series (data, name of the time series and dates). For the series marked in the multi-document **Copy** copies the specification details. It can be pasted into external application. Similarly, **Copy series** copies the time series, which can be then pasted into another application. Both **Cut** and **Delete** remove the marked time series from the multi-document.

- **Clear selection** – unmarks the series in the SAProcessing window. Option is inactive when none of the time series in the SAProcessing window are marked.

- **Specification** – enables the user to change the specification used for seasonal adjustment of a selected series to one chosen from the list of the pre-defined and user-defined specifications.

- **Priority** – indicator that can be used to mark series that require more or less attention. The Priority parameter takes values from 0 to 10. JDemetra+ computes it automatically, based on the average of the (logged) series. The user can chose the method of computation (logbased or level based).

- **Initial Order** – displays times series on the list in initial order. The option restores the initial order if the list has been sorted by given column (e.g. by quality or by method).

- **Output** – enables the user to export the results to a given output formats (TXT, XLS, ODBC, CSV, CSV matrix), to specify the destination folder and to customise the content of the exported file. The list of available output items is determined by the chosen output format and it can be found in 7.7. The **Output** option is discussed in detail in the ‘JDemetra+ User Guide’ (2016), item 3.2.2.1.

- **Report** – generates a summary report from the multi-process, that includes information about e.g. number of series, specifications used, models summary, diagnostic summary. An example of the output is presented below.
Figure 3.48: The report from the multi-process.
4. Modelling

4.1. Specifications

Specifications described in this section are sets of parameters and values assigned to them that contain all information necessary for time series modelling. The default critical values used by the tests included in the specifications can be changed by the user in the Tools $\rightarrow$ Options menu (see 3.4.7).

The Specifications section, which is a part of the Workspace window, contains a set of pre-defined specifications that enables the user to model the time series using two options: the TRAMO model or the RegARIMA model. The set of pre-defined modelling specifications is presented in Table 4.1. It contains the most commonly used specifications for seasonal adjustment. Pre-defined specifications correspond to the terminology used in TSW+. The users are strongly recommended to start their analysis with one of those specifications (usually RG4c or RG5c for RegARIMA and TR4 or TR5 for TRAMO) then, if need be, to change some of the options afterwards using the Specification button (see the ‘JDemetra+ User Guide’ (2016), 3.3.2). The default specification for TRAMO is TR5, while for RegARIMA it is RG4c.

Table 4.1: Pre-defined modelling specifications.

<table>
<thead>
<tr>
<th>Settings</th>
<th>Transformation</th>
<th>Pre-adjustment for leap-year</th>
<th>Working days</th>
<th>Trading days</th>
<th>Easter effect</th>
<th>Outliers</th>
<th>ARIMA model</th>
</tr>
</thead>
<tbody>
<tr>
<td>TR0</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>(0,1,1)(0,1,1)</td>
</tr>
<tr>
<td>TR1</td>
<td>test</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>test</td>
<td>(0,1,1)(0,1,1)</td>
</tr>
<tr>
<td>TR2</td>
<td>test</td>
<td>no</td>
<td>test</td>
<td>no</td>
<td>test</td>
<td>test</td>
<td>(0,1,1)(0,1,1)</td>
</tr>
<tr>
<td>TR3</td>
<td>test</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>test</td>
<td>AMI</td>
</tr>
<tr>
<td>TR4</td>
<td>test</td>
<td>no</td>
<td>test</td>
<td>no</td>
<td>test</td>
<td>test</td>
<td>AMI</td>
</tr>
<tr>
<td>TR5</td>
<td>test</td>
<td>no</td>
<td>no</td>
<td>test</td>
<td>test</td>
<td>test</td>
<td>AMI</td>
</tr>
<tr>
<td>RG0</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>(0,1,1)(0,1,1)</td>
</tr>
<tr>
<td>RG1</td>
<td>test</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>test</td>
<td>(0,1,1)(0,1,1)</td>
</tr>
<tr>
<td>RG2c</td>
<td>test</td>
<td>test</td>
<td>test</td>
<td>no</td>
<td>test</td>
<td>test</td>
<td>AMI</td>
</tr>
<tr>
<td>RG3</td>
<td>test</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>test</td>
<td>AMI</td>
</tr>
<tr>
<td>RG4c</td>
<td>test</td>
<td>test</td>
<td>test</td>
<td>no</td>
<td>test</td>
<td>test</td>
<td>AMI</td>
</tr>
<tr>
<td>RG5c</td>
<td>test</td>
<td>test</td>
<td>no</td>
<td>test</td>
<td>test</td>
<td>test</td>
<td>AMI</td>
</tr>
</tbody>
</table>

Explanations for settings:

---

13 For a description of both models see 7.2.1.

- **Transformation test** – a test is performed to choose between additive decomposition (no transformation) and multiplicative decomposition (logarithmic transformation).

- **Pre-adjustment for leap-year** – a correction of the February values applied to the original series before logarithmic transformation. The original values in February are multiplied by \( \frac{28.25}{29} \) for leap years and by \( \frac{28.25}{28} \) for non-leap years. Values for other months are not modified.

- **Working days** – a pre-test is made for the presence of the working day effect by using one parameter specification.

- **Trading days** – a pre-test is made for the presence of the trading day effect by using six parameters specification.

- **Easter effect** – the tests for the necessity of a correction for the Easter effect in the original series. The length of the Easter effect, which is considered here, is 6 days (for the TRAMO specifications) and 8 days (for the RegARIMA specifications).

- **Outliers** – An automatic identification of three types of outlier: AO (additive outliers), LS (level shifts), TC (transitory changes) using default critical values.

- **ARIMA model** – the choice between fixing the ARIMA model structure to \( (0,1,1)(0,1,1) \) or searching for the ARIMA model using automatic model identification procedure (AMI). The \( (0,1,1)(0,1,1) \) model (called the Airline model) is used as a default model in several TRAMO/SEATS and X-13ARIMA-SEATS specifications because it has been shown in many studies that this model is appropriate for many real monthly or quarterly time series with a statistically significant seasonal pattern. Moreover, the Airline model approximates well many other models and provides an excellent "benchmark" model\(^{15}\).

The user may add new modelling specifications to the *Workspace* window. To do this, go to the *Modelling* section, right click on the *tramo* or *regarima* item in the *specifications* node and select *New* from the local menu.

![Figure 4.1: Creating a new specification in the Modelling section.](image)

---

\(^{15}\) MARAVALL, A. (2009).
4.1.1. Tramo

This section discusses the options available for the TRAMO specifications, which are based on the TSW+ program developed by Agustin Maravall and Victor Gómez. It is divided into five parts that correspond to the TRAMO specification sections and are presented in the order in which they are displayed in the graphical interface of JDemetra+.

![Figure 4.2: The TRAMO specification sections.](image)

To facilitate the comparison between JDemetra+ specifications and specifications used in TSW+, under each option the name of the corresponding specification and argument from the original software is given; small variants are indicated by an asterisk in the tables in 4.1 and 4.5 For an exact description of the different parameters, the user should refer to the documentation of TSW+. Some additional explanations about the TRAMO model, its parameters and estimation procedure are given in 7.1.1.

For the pre-defined specifications the parameters are disabled, while in the case of the user-defined specification the user can set them individually. However, as in some cases the choice of a given value results in limitation of the possible alternatives for other parameters, the user is not entirely free to set the parameters values. Most arguments have default values; these are given in the documentation and used unless changed by the user.

4.1.1.1. Estimate

The Estimate section specifies the details of estimation procedure of the TRAMO model determined in the Regression and Arima sections, which are explained in the items 4.1.1.3 and 4.1.1.5.

Table 4.2: TRAMO specification – options for the Estimate section.
<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
</table>
| Model span → type      | Specifies the span (data interval) of the time series to be used for the estimation of the RegARIMA model coefficients. The RegARIMA model is then applied to the whole series. With this argument the data early in the series can be prevented from affecting the forecasts, or, alternatively, data late in the series are excluded from the modelling span, so that they do not influence the regression estimates. The available parameter’s values are:  
  - All – full time series span is considered in the modelling;  
  - From – date of the first time series observation included in the pre-processing model;  
  - To – date of the last time series observation included in the pre-processing model;  
  - Between – date of the first and the last time series observations included in the pre-processing model;  
  - Last – number of observations from the end of the time series included in the pre-processing model;  
  - First – number of observations from the beginning of the time series included in the pre-processing model;  
  - Excluding – number of observations excluded from the beginning (specified in the first field) and/or end (specified in the last field) of the time series in the pre-processing model. With the options Last, First, Excluding the span can be computed dynamically on the series. The default setting is All. |
| Tolerance              | Convergence tolerance for the nonlinear estimation. The absolute changes in the log-likelihood are compared to Tolerance to check the convergence of the estimation iterations. The default setting is 0.0000001. |
| Exact ML               | When this option is marked, an exact maximum likelihood estimation is performed. Alternatively, the Unconditional Least Squares method is used. However, in the current version of JDemetra+ it is not recommended to change this parameter’s value. |
| Unit root limit        | Limit for the autoregressive roots. If the inverse of a real root of the autoregressive polynomial of the ARIMA model is higher than that limit, the root is set equal to 1. The default parameter value is 0.96. |

### 4.1.1.2. Transformation

The Transformation section is used to transform the series prior to estimating the TRAMO model.
Table 4.3: TRAMO specification – options for the Transformation section.

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Function</td>
<td>Transformation of data. The user can choose between:</td>
</tr>
<tr>
<td>Transformation;</td>
<td>- None – no transformation of data;</td>
</tr>
<tr>
<td>lam</td>
<td>- Auto – the program tests for the log-level specification. This option is</td>
</tr>
<tr>
<td></td>
<td>recommended for automatic modelling of many series.</td>
</tr>
<tr>
<td></td>
<td>- Log – takes logs of data.</td>
</tr>
<tr>
<td></td>
<td>The default setting is Auto.</td>
</tr>
<tr>
<td>Fct</td>
<td>Control of the bias in the log/level pre-test (the function is active</td>
</tr>
<tr>
<td>Transformation;</td>
<td>when Function is set to Auto); Fct &gt; 1 favours levels, Fct &lt; 1 favours logs.</td>
</tr>
<tr>
<td>fct</td>
<td>The default setting is 0.95.</td>
</tr>
</tbody>
</table>

4.1.1.3. Regression

The Regression section allows for estimating deterministic effects in the series with the pre-defined regression variables. These variables include user-defined variables, several types of pre-specified outliers, intervention variables and calendar effects. The pre-defined and user-defined regression variables are selected with the variables argument.

Table 4.4: TRAMO specification – options for the Regression section.

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calendar →</td>
<td>Determines the manner in which the calendar effects are entered</td>
</tr>
<tr>
<td>tradingDays →</td>
<td>in the TRAMO model. It can be done in two ways: automatically</td>
</tr>
<tr>
<td>automatic</td>
<td>(on a basis of the specified test) or manually. The calendar effects</td>
</tr>
<tr>
<td></td>
<td>that are considered here are: trading day, working day and leap</td>
</tr>
<tr>
<td></td>
<td>year. The significance of the Easter effect is considered in the</td>
</tr>
<tr>
<td></td>
<td>following part of this section.</td>
</tr>
<tr>
<td></td>
<td>- Unused – the calendar effects included in the TRAMO model are those</td>
</tr>
<tr>
<td></td>
<td>specified by the user through the Option,</td>
</tr>
<tr>
<td></td>
<td>tradingDays and LeapYear parameters.</td>
</tr>
</tbody>
</table>

---

16 The test for log-level specification used by TRAMO is based on the maximum likelihood estimation of the parameter $\lambda$ in the Box-Cox transformation, which is a power transformation such that the transformed values of the time series $y$ are a monotonic function of the observations, i.e. $y^\gamma = \begin{cases} y^{\lambda} & \lambda \neq 0 \\ y & \lambda = 0 \end{cases}$. The program first fits two Airline models (i.e. ARIMA (0,1,1)(0,1,1)) with log $y^\gamma = \cdots$
mean to the time series: one in logs ($\lambda = 0$), other without logs ($\lambda = 1$). The test compares the sum of squares of the model without logs with the sum of squares multiplied by the square of the geometric mean of the (regularly and seasonally) differenced series in the case of the model in logs. Logs are taken in the case this last function is the minimum. GÓMEZ, V., and MARAVALL, A. (2010).

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>▪ FTest – the choice of the number of calendar variables to be included in the TRAMO model is based on the outcome of Ftests computed on the models with the trading day variables and with the working day variable (F-test is performed on the coefficients of the calendar variables). The model with higher F value is chosen, provided that it is higher than Pftd (see Calendar → tradingDays → Pftd).</td>
<td></td>
</tr>
<tr>
<td>▪ WaldTest – working days are restricted to trading days. The choice of the number of calendar variables to be included in the TRAMO model is based on the outcome of the Wald test computed on that restriction. The F-test computed on the preferred model must be higher than Pftd (see Calendar → tradingDays → Pftd).</td>
<td></td>
</tr>
</tbody>
</table>

The default setting is Unused.

Calendar → tradingDays → Pftd
Trading day; pftd

P-Value applied in the test specified by the automatic parameter to assess the significance of the pre-tested calendar effect and to decide if calendar effects are included into the TRAMO model. The Pftd option is displayed when Calendar → tradingDays → automatic is set to FTest or WaldTest. The default Pftd setting is 0.01.
Specifies the type of calendar being assigned to the series. The following types of calendar estimation are available:

- **None** – means that calendar effects will not be included in the regression.
- **Default** – means that a default JDemetra+ calendar, which does not include any country-specific holidays, will be used.
- **Stock** – estimates day-of-week effects for inventories and other stock reported for the $w^{th}$ day of the month (to denote the last day of the month enter 31).
- **Holidays** – corresponds to the pre-defined trading day variables, modified to take into account country specific holidays. It means that after choosing this option the user should specify the type of trading days effect (Calendar → tradingDays → TradingDays) and a previously defined calendar that includes the country specific holidays (Calendar → tradingDays → holidays).
- **UserDefined** – used when the user wants to specify his own trading day variables. With this option calendar effects are captured only by regression variables chosen by the user from the previously created list of user-defined variables (see 3.1.2).

The default setting is Default.

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Calendar → tradingDays → option</strong></td>
<td>Specifies type of calendar being assigned to the series.</td>
</tr>
<tr>
<td><strong>Trading day; itr</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Calendar → tradingDays → holidays</strong></td>
<td>Regression variables; ireg*</td>
</tr>
<tr>
<td><strong>Regression variables; ireg</strong></td>
<td>List of user-defined calendars to be used to create calendar regression variables. Option is available when Calendar → tradingDays → option is set to Holidays. The user is expected to click the field to expand a list of previously defined calendars and choose an appropriate item. The default setting is Default, which implies that the default calendar is used and no country-specific holidays are considered.</td>
</tr>
<tr>
<td><strong>Calendar → tradingDays → userVariables</strong></td>
<td>Regression variables; ireg, user, usertype= (...td...)*</td>
</tr>
<tr>
<td><strong>Regression variables; ireg, user, usertype= (...td...)</strong></td>
<td>List of predefined regression variables to be included in the model. Option is available when Calendar → tradingDays → option is set to UserDefined. When the user chooses the UserDefined type for the trading day effect estimation, one must specify the corresponding variables by clicking the field and choosing variables from the list. It should be noted that such variables must have been previously defined (see 3.1.2), otherwise the list is empty. The default setting is Unused.</td>
</tr>
</tbody>
</table>
### Calendar → tradingDays → tradingDays

**regression variables, itrad**

Assigns a type of model-estimated regression effect to pre-specified regression variables. Option is available when **Calendar → tradingDays → automatic** is set to **Unused** and **Calendar → tradingDays → option** is set to **Default** or **Holidays**. Acceptable values:

- **None** – excludes calendar variables from the regression model.
- **TradingDays** – includes six day-of-the-week regression variables.
- **WorkingDays** – includes the working/non-working day contrast variable.

The default setting is **TradingDays**. Some options can be disabled when the **Adjust** option is used (see 4.2.3).

### Calendar → tradingDays → leapYear

**regression variables; itrad**

Enables/disables a leap-year correction. By default, the checkbox is marked which implies that the leap-year effect correction is enabled.

### Calendar → tradingDays → RegressionTestType

**regression variables; itrad**

Option for pre-test of trading day effects. It is not available when the **Calendar → tradingDays → automatic** checkbox is marked.

- **None** – the test is not performed; the specified calendar variables are used in the model without pre-testing.
- **Separate_T** – a t-test is applied to each trading day variable separately. The trading day variables are included in the TRAMO model if at least one t-statistic is greater than 2.6 or if two t-statistics are greater than 2.0 (in the absolute values).
- **Joint_F** – a joint F-test of significance of all the trading day variables. The trading day effect is significant if the F statistic is greater than 0.95.

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>The default setting is <strong>Separate_T</strong>.</td>
</tr>
</tbody>
</table>
### Calendar → easter → Option  
\textit{easter effect; ieast}\textsuperscript{*}

Option for specification the presence and length of the Easter effect.
- \textit{Unused} – the Easter effect is not considered;
- \textit{Standard} – the Easter effect influences the period of $n$ days strictly before Easter Sunday where $n$ is \textbf{Duration} parameter value;
- \textit{Include Easter} – the Easter effect influences the entire period ($n$) up to and including Easter Sunday, where $n$ is \textbf{Duration} parameter value;
- \textit{Include Easter Monday} – the Easter effect influences the entire period ($n$) up to and including Easter Monday, where $n$ is \textbf{Duration} parameter value.

The default setting is \textit{Standard}.

### Calendar → easter → Use Julian Easter  
\textemdash; \textemdash;

When marked, it enables to include Easter, which date derives from Julian calendar and is converted to Gregorian calendar date.

By default, the checkbox is unmarked.

### Calendar → easter → duration  
\textit{easter effect; idur}

Duration (length in days) of the Easter effect. The length of the Easter effect can range from 1 to 15 days. The default value is 6.

### Calendar → easter → test  
\textit{easter effect; ieast}

A t-test applied for the significance of the Easter effect. The Easter effect is considered as significant if the t-statistic is higher than 1.96.

By default, the checkbox is marked which implies that the test is used.

### Pre-specified outliers  
\textit{regression variables; –}

User-defined outliers are used when prior knowledge suggests that certain effects exist at known time points\textsuperscript{16}. Four pre-defined outlier types, which are simple forms of intervention variables, are implemented:
- Additive Outlier (AO);
- Level shift (LS);
- Temporary change\textsuperscript{17} (TC);
- Seasonal outliers (SO).

Descriptions and formulas are available in 7.1.1.

By default, no pre-specified outliers are included in the specification.

### Intervention variables  
\textit{regression variables; –}

The intervention variables are defined as in TSW+. Following the definition, these effects are special events known a-priori (strikes, devaluations, political events, and so on). Intervention variables are modelled as any possible sequence of ones and zeros, on


\textsuperscript{17} In the TRAMO/SEATS method this type of outlier is called transitory change.
<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
</table>
|                 | which some operators may be applied. Intervention variables are built as combinations of five basic structures:\18:
|                 | ▪ Dummy variables\19;                                                                                                                         |
|                 | ▪ Any possible sequence of ones and zeros;                                                                                                  |
|                 | ▪ \[\frac{1}{(1-\delta s B^s)} (0 < \delta \leq 1)\];                                                                                     |
|                 | ▪ \[\frac{1}{(1-\delta s B^s)} (0 < \delta \leq 1)\];                                                                                     |
|                 | Where B is backshift operator (i.e. \(B^k X_t = X_{t-k}\)) and s is frequency of the time series (s = 12 for a monthly time series, s = 4 for a quarterly time series). |
|                 | These operations enable the generation of not only AO, LS, TC, SO and RP outliers but also sophisticated intervention variables that are well-adjusted to the particular case. By default, no intervention variables are included in the specification. |

### Ramps

**Regression variables:** –

A ramp effect means a linear increase or decrease in the level of the series over a specified time interval \(t_0\) to \(t_1\). All dates of the ramps must occur within the time series span. Ramps can overlap other ramps, additive outliers and level shifts. The graph and formula are available in the 7.1.1. By default, no ramps are included in the specification.

---


\(^{19}\) Dummy variable is the variable that takes the values 0 or 1 to indicate the absence or presence of some effect.
The user-defined variable is an external regressor included by the user to the TRAMO model. To add a user-defined variable to the model, one must specify the corresponding variable by clicking the Name field and choosing a variable from the list. It should be noted that such variables must have been previously defined (see 3.1.2), otherwise the list is empty. The user-defined regression variable associated to a specific component should not contain effects that have to be associated with another component. Therefore, the following rules should be obeyed:

- The variable assigned to the trend or to the seasonally adjusted series should not contain a seasonal pattern;
- The variable assigned to the seasonal should not contain a trend (or level);
- The variable assigned to the irregular should contain neither a seasonal pattern nor a trend (or level).

The user-defined variables should cover the period of the dependent series and the appropriate number of forecasts. For the other periods the variables are implicitly set to 0. If the forecasts are not provided, it will not alter the results of the seasonal adjustment but the forecasts of the final components will be unusable.

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
</table>

User-defined variables
regression variables; –
The effect of the user-defined variable can be assigned to the:
- Trend;
- Irregular component;
- Seasonal component;
- Seasonally adjusted series;
- None of the above, which implies that it exists as an additional component (option *Undefined*), which is a default setting. With this option the regression variable is used to improve the modelling, but it is not removed from the series for the decomposition.

The calendar component is not available in this section. Therefore, a user-defined variable assigned to the calendar effect should be added in the calendar part of the specification. For the user-defined variable the structure of the lags can be specified using the options *first lag* and *last lag*. When the regression variable $x_t$ is introduced with *first lag* $= l_a$ and *last lag* $= 1$, JDemetra+ includes in the TRAMO model a set of variables, $x_{t-0}, x_{t-1}, \ldots, x_{t-l_b}$, and estimates the respective regression coefficients called the impulse response weights.

To include only the first lag ($x_{t-1}$) of the user-defined variable ($x_t$) the user should put *first lag* $= 1$. If for a monthly series one puts *first lag* $= 0$ and *last lag* $= 11$, it means that in addition to the instantaneous effect of the user-defined variable, also the effects of 11 lagged explanatory variables are included in the model. In this case the set of estimated coefficients, called a transfer function, describe how the changes in $x_t$ that took place over a year are transferred to the dependent variable. However, the lagged variables are often collinear so that caution is needed in attributing much meaning to each coefficient.

By default, no user-defined variables are included in the specification.

---

20. The convention for assigning effects for user-defined regressors is based on the TRAMO/SEATS solution. The original X-12ARIMA program does not provide the same functionality. The effect of additional component (option *Undefined*) is a special feature introduced in the original X-12 ARIMA program and not available in TRAMO/SEATS.

21. A typical example could be a variable that measures the weather conditions: introducing them could improve the linearization; however, these effects are not allocated to a specific component. On the contrary, the impact of this variable will be split between all components.

22. More details and examples can be found in MARAVALL, A. (2008).

4.1.1.4. Outliers

The Outliers section enables the user to perform an automatic detection of additive outliers, temporary change outliers, level shifts, seasonal outliers, or any combination of the four, using the specified model.

Table 4.5: TRAMO specification – options for the Outliers section.

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>IsEnabled &lt;i&gt;outliers; iatip&lt;/i&gt;</td>
<td>Enables/disables the automatic detection of outliers in the span determined by the Detection span option. By default, the checkbox is marked which implies that the automatic identification of outliers is enabled.</td>
</tr>
<tr>
<td>Use default critical value &lt;i&gt;outliers; va&lt;/i&gt;</td>
<td>The critical value is automatically determined by the number of observations in the interval specified by the Detection span option. When Use default critical value is disabled, the procedure uses the critical value inputted in the Critical value item (see below). Otherwise, the default value is used (the first case corresponds to &quot;critical = xxx&quot;; the second corresponds to a specification without the critical argument). It should be noted that it is not possible to define a separate critical value for each outlier type. By default, the checkbox is marked which implies that the automatic determination of the critical value is enabled.</td>
</tr>
<tr>
<td>Critical value &lt;i&gt;outliers; va&lt;/i&gt;</td>
<td>The critical value used in the outlier detection procedure. The option is active once Use default critical value is disabled. By default, it is set to 3.5.</td>
</tr>
</tbody>
</table>
## Detection span → type

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Additive outliers; aio*</td>
<td>Automatic identification of additive outliers. By default, this option is enabled.</td>
</tr>
<tr>
<td>Level shift outliers; aio*</td>
<td>Automatic identification of level shifts. By default, this option is enabled.</td>
</tr>
<tr>
<td>Transitory outliers; aio*</td>
<td>Automatic identification of transitory changes. By default, this option is enabled.</td>
</tr>
<tr>
<td>Seasonal outliers; aio*</td>
<td>Automatic identification of seasonal outliers. By default, this option is disabled.</td>
</tr>
<tr>
<td>EML estimation outliers; imvx</td>
<td>The estimation method used in the automatic model identification procedure. By default, the fast method of Hannan-Rissanen is used for parameter estimation in the intermediate steps of the automatic detection and correction of outliers. When the checkbox is marked the exact maximum likelihood estimation method is used.</td>
</tr>
<tr>
<td>TC rate outliers; deltatc</td>
<td>The rate of decay for the transitory change outlier. It takes values between 0 and 1. The default value is 0.7.</td>
</tr>
</tbody>
</table>

### 4.1.1.5. Arima

Identification of the ARIMA part of the RegARIMA model can be done either in an automatic way or by the user who specifies the appropriate parameters. This choice is controlled by the **Automatic**
option and results in a list of parameters specific to the chosen ARIMA identification procedure. When the **Automatic** option is marked, the order of the ARIMA model results from the automatic identification procedure. The maximum order of the regular polynomials is 3, and the maximum order of seasonal polynomials is 1. The parameters available for automatic model identification are presented below.

Table 4.6: TRAMO specification – options for the automatic identification of the ARIMA model.

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Automatic</strong></td>
<td>When marked it enables automatic modelling of the ARIMA model to be performed.</td>
</tr>
<tr>
<td><code>automdl; ami; idif, inic*</code></td>
<td>5. A limit for the AR and the MA roots to be assumed equal(^{24}). This option is used in the automatic identification of the differencing order. If the difference in moduli of an AR and an MA root (when estimating ARIMA(1,0,1)(1,0,1) models in the second step of the automatic identification of the differencing polynomial) is smaller than <strong>Cancellation limit</strong>, the two roots cancel out. The default parameter value is 0.05.</td>
</tr>
<tr>
<td><strong>Cancelation limit</strong></td>
<td>The threshold value for the initial unit root(^{25}) test in the automatic differencing procedure. When one of the roots in the estimation of the (2,0,0)(1,0,0) plus mean model, which is performed in the first step of the automatic model identification procedure, is larger than <strong>First unit root limit</strong>, in modulus, it is set equal to unity. This value should be less than 1 and greater than 0.8. The default parameter value is 0.97.</td>
</tr>
</tbody>
</table>

\(^{24}\) Cancellation issue is described in 7.1.1.6.

\(^{25}\) A unit root is an attribute of a statistical model of a time series whose autoregressive parameter is one.
Final UR (Diff.)

The threshold value for the final unit root test in the automatic differencing procedure. When one of the roots in the estimation of the $(1,0,1)(1,0,1)$ plus mean model, which is performed in the second step of the automatic model identification procedure, is larger than Second unit root limit, in modulus, it is checked if there is a common factor in the corresponding AR and MA polynomials of the ARMA model that can be cancelled (see Cancelation limit). If there is no cancellation, the AR root it is set equal to unity (i.e. the differencing order changes). The value of the Second unit root limit should be less than 1 and greater than 0.8. The default parameter value is 0.91.

Reduce CV

The percentage by which the outlier critical value will be reduced when the preferred model is found to have a Ljung-Box Q-statistic with an unacceptable confidence coefficient (7.1.1.1). The parameter should be between 0 and 1, and will only be active when automatic outlier identification is selected. The reduced critical value will be set to $(1-\text{ReduceCV})\times CV$, where $CV$ is the original critical value. The default parameter value is 0.12.

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ArmaLimit</td>
<td>The threshold value for t-statistics of ARMA coefficients used for the final test of model parsimony. If the highest order ARMA coefficient has a t-value less than this value in magnitude, JDemetra+ will reduce the order of the model. The value given for ArmaLimit is also used for the final check of the constant term; if the constant term has a t-value less than ArmaLimit in magnitude, the program will remove the constant term from the set of regressors. The ArmaLimit value should be greater than zero. The default parameter value is 1.</td>
</tr>
<tr>
<td>LjungBox limit</td>
<td>Acceptance criterion for the confidence intervals of the Ljung-Box Q-statistic. If the Ljung-Box Q-statistics for the residuals of a final model (checked at lag 24 if the series is monthly, 16 if the series is quarterly) is greater than LjungBox limit, the model is rejected, the outlier critical value is reduced, and model and outlier identification (if specified) is redone with a reduced value (see the Reduce CV argument). After two unfruitful attempts, a default</td>
</tr>
</tbody>
</table>
A model (usually (3,1,1)(0,1,1)) will be used. The default parameter value is 0.95.

<table>
<thead>
<tr>
<th><strong>Accept Default</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Controls whether the default model (ARIMA(0,1,1)(0,1,1)) may be chosen in the first step of the automatic model identification. More explicitly, if the Ljung-Box Q-statistics for the residuals is acceptable, the default model is accepted and no further attempt will be made to identify and other. By default, the <strong>Accept Default</strong> option is false.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Compare to default</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>6. If marked, it compares the model identified by the automatic procedure to the default model (ARIMA(0,1,1)(0,1,1)) and the model with the best fit is selected. The criteria for comparison are the residual diagnostics from the automatically identified model and those of the default model (the residual standard error and the confidence coefficient of the Ljung-Box Q-statistic), the number of outliers and the structure and estimated parameters of the model identified by the automatic procedure. The comparison is done because the default model is robust and departure from this model can be unstable. By default, the <strong>Compare to default</strong> option is false.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Balanced</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Controls whether the automatic model procedure will have a preference for balanced model (i.e., models for which the order of the combined AR and differencing operators is equal to the order of the combined MA operator 28). By default, the <strong>Balanced</strong> checkbox is unmarked. When it is checked the same preference as the TRAMO program is used.</td>
</tr>
</tbody>
</table>

When the **Automatic** checkbox in the **Arima** section is unmarked, JDemetra+ allows the user to specify the structure of the ARIMA part of the RegARIMA model. Initial values for the individual AR and MA parameters can be specified for the iterative estimation. Also, individual parameters can be held fixed at these initial values while the rest of the parameters are estimated. The options available here correspond to the original X-13ARIMA-SEATS **arima** spec with some limitations. JDemetra+ does not allow for operators with missing lags. Also the maximum lag is reduced in comparison with Win X-13.

---

Individual parameters can be held fixed at these initial values while the rest of the parameters are estimated. However, users should not specify initial values for the MA parameters that yield the MA polynomial with roots inside the unit circle.

Table 4.7: TRAMO specification – options for manual identification of the ARIMA model.

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Automatic am; idif;inic</td>
<td>When unmarked it enables the user to enter the parameters of the ARIMA model.</td>
</tr>
<tr>
<td>Mean mean; imean</td>
<td>When marked it is considered that the mean is part of the ARIMA model (it highly depends on the chosen model).</td>
</tr>
<tr>
<td>P arima; p</td>
<td>The order of the non-seasonal autoregressive polynomial. The maximum order of the non-seasonal autoregressive polynomial is 4. The default value is 0.</td>
</tr>
</tbody>
</table>
| phi arima; phi, jpr | Coefficients of the non-seasonal, autoregressive polynomial (AR). If used, each non-seasonal AR parameter in the model requires a label that indicates the procedure of its estimation. The choice can be made from:  
  ▪ Undefined – estimates a parameter without the use of any user defined input (the default value).  
  ▪ Initial – estimates a parameter using as initial condition the value defined by the user.  
  ▪ Fixed – holds a parameter fixed during estimation at the value defined by the user. |
| D arima; d | Non-seasonal differencing order. The maximum number of nonseasonal differences is 2. The default value is 1. |
| Q arima; q | The order of the non-seasonal moving average polynomial. The maximum order of the non-seasonal moving average polynomial is 4. The default value is 1. |
| theta arima; th, jqr | Coefficients of the parameters of the non-seasonal, moving average polynomial (MA). If used, to each non-seasonal MA parameter in the model a label that indicates the procedure of its estimation should be assigned. The choice can be made from:  
  ▪ Undefined – estimates a parameter without the use of any user defined input (the default value). |
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BP (\text{arima;bp})</td>
<td>The order of the seasonal autoregressive polynomial. The default value is 0.</td>
</tr>
</tbody>
</table>
| Bphi \(\text{arima;bphi, jps}\) | Coefficients of the seasonal autoregressive polynomial (AR). If used, to each seasonal AR parameter in the model a label that indicates the procedure of its estimation should be assigned. The choice can be made from:  
  - **Undefined** – estimates a parameter without the use of any user defined input (the default value).  
  - **Initial** – estimates a parameter using as initial condition the value defined by the user.  
  - **Fixed** – holds a parameter fixed during estimation at the value defined by the user. |
| BD \(\text{arima;bd}\) | Seasonal differencing order. The maximum number of seasonal differences is 1. The default value is 1. |
| BQ \(\text{arima;bq}\) | The order of the seasonal moving average polynomial. The maximum order of the seasonal moving average polynomial is 1. The default value is 1. |
| Btheta \(\text{arima; bth, jqs}\) | Coefficients of the parameters of the seasonal moving average polynomial (MA). If used, each seasonal MA parameter in the model requires a label that indicates the procedure of its estimation. The choice can be made from:  
  - **Undefined** – estimates a parameter without the use of any user defined input (the default value).  
  - **Initial** – estimates a parameter using as initial condition the value defined by the user.  
  - **Fixed** – holds a parameter fixed during estimation at the value defined by the user. |

### 6.1.1. RegARIMA

This section discusses the options available for the RegARIMA specifications, which are based on the original X-13ARIMA-SEATS program developed by the U.S. Census Bureau. The RegARIMA specifications are - to a very large extent - organised according to the different individual specifications of the original program and are presented in the order in which they are displayed in the graphical interface of JDemetra+. 
To facilitate the comparison between JDemetra+ specifications and specifications used in Win X13, under each option the name of the corresponding specification and argument from the original software is given, if any. For an exact description of the different parameters, the user should refer to the documentation of the original X-13ARIMA-SEATS program\textsuperscript{29}. Some additional explanations about the RegARIMA model, its parameters and estimation procedure are given in the Annex.

For pre-defined specifications the parameters are fixed, while in the case of the user-defined specification the user can set them individually. However, as in some cases the choice of a given value results in limitation of the possible alternatives for other parameters, the user is not entirely free to set the parameters values. Most arguments have default values; these are given in the documentation and used unless changed by the user.

### 6.1.1.1. Estimate

The *Estimate* section specifies the details of estimation procedure of the RegARIMA model determined in the *Regression* and *Arima* sections, which are explained in the items 4.1.2.3 and 4.1.2.4.

#### Table 4.8: RegARIMA specification – options for the Estimate section.

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model span → type</strong></td>
<td>Specifies the span (data interval) of the time series to be used for the estimation of the RegARIMA model coefficients. The RegARIMA model is then applied to the whole series. With this argument the data early in the series can be prevented from affecting the forecasts, or, alternatively, data later in the series are excluded from the modelling span, so that they do not influence the regression estimates. The available parameter values are:</td>
</tr>
</tbody>
</table>

### All – means that full time series span is considered in the modelling:

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>From – date of the first time series observation included in the pre-processing model;</td>
<td></td>
</tr>
<tr>
<td>To – date of the last time series observation included in the pre-processing model;</td>
<td></td>
</tr>
<tr>
<td>Between – date of the first and the last time series observations included in the pre-processing model;</td>
<td></td>
</tr>
<tr>
<td>Last – number of observations from the end of the time series included in the pre-processing model;</td>
<td></td>
</tr>
<tr>
<td>First – number of observations from the beginning of the time series included in the pre-processing model;</td>
<td></td>
</tr>
<tr>
<td>Excluding – number of observations excluded from the beginning (specified in the first field) and/or end (specified in the last field) of the time series in the pre-processing model. With the options Last, First, Excluding the span can be computed dynamically on the series. The default setting is All.</td>
<td></td>
</tr>
</tbody>
</table>

**Tolerance estimate{tol}**

Convergence tolerance for the nonlinear estimation. The absolute changes in the log-likelihood function are compared to Tolerance to check for the convergence of the estimation iterations. The default setting is 0.0000001.

### 6.1.1.2. Transformation

The Transformation section is used to transform the series prior to estimating the RegARIMA model.

**Table 4.9: RegARIMA specification – options for the Transformation section.**

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Function transform{function=}</td>
<td>Transformation of data. The user can choose between:</td>
</tr>
<tr>
<td></td>
<td>- None – no transformation of data;</td>
</tr>
<tr>
<td></td>
<td>- Log – takes logs of data;</td>
</tr>
<tr>
<td></td>
<td>- Auto – the program tests for the log-level specification. This option is recommended for automatic modelling of many series.</td>
</tr>
</tbody>
</table>
The default setting is *Auto*.

---

The test for log-level specification used by TRAMO/SEATS is based on the maximum likelihood estimation of the parameter $\lambda$ in the Box-Cox transformations (which is a power transformations such that the transformed values of the time series $y$ are a monotonic function of the observations, i.e. $y^\tau = \left\{ \begin{array}{ll} y^\tau & \lambda \neq 0 \\ \lambda \end{array} \right. = 0$). The program first fits two Airline models (i.e. ARIMA log $y_i$ (0,1,1)(0,1,1)) to the time series: one in logs ($\lambda = 0$), other without logs ($\lambda = 1$). The test compares the sum of squares of the model without logs with the sum of squares multiplied by the square of the geometric mean in the case of the model in logs. Logs are taken in the case this last function is the maximum, GÓMEZ, V., and MARAVALL, A. (2009).

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIC difference</td>
<td>Defines the difference in AICC needed to accept no transformation over a log transformation when the automatic transformation selection option is invoked. The option is disabled when <strong>Function</strong> is not set to <em>Auto</em>. The default AIC difference value is -2.</td>
</tr>
<tr>
<td>Adjust transform</td>
<td>Options for proportional adjustment for the leap year effect. The option is available when <strong>Function</strong> is set to <em>Log</em>. <strong>Adjust</strong> can be set to:</td>
</tr>
<tr>
<td></td>
<td>▪ <em>LeapYear</em> – performs leap year adjustment of monthly or quarterly data;</td>
</tr>
<tr>
<td></td>
<td>▪ <em>LengthofPeriod</em> – performs length-of-month adjustment on monthly data or length-of-quarter adjustment on quarterly data;</td>
</tr>
<tr>
<td></td>
<td>▪ <em>None</em> – does not include a correction for the length of the period. The default setting is <em>None</em>.</td>
</tr>
</tbody>
</table>

### 6.1.1.3. Regression

The **Regression** section allows for estimating deterministic effects in the series with the pre-defined regression variables. These variables include user-defined variables, several types of pre-specified outliers, intervention variables and calendar effects. The pre-defined and user-defined regression variables are selected with the variables argument.

**Table 4.10: RegARIMA specification – options for the Regression section.**
<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calendar → tradingDays → option</td>
<td>Specifies the type of calendar being assigned to the series. The following types of calendar estimation are available:</td>
</tr>
<tr>
<td></td>
<td>▪ None – excludes calendar variables from the regression model.</td>
</tr>
<tr>
<td></td>
<td>▪ Default – means that a default JDemetra+ calendar, which does not include any country-specific holidays, will be used.</td>
</tr>
<tr>
<td></td>
<td>▪ Stock – estimates day-of-week effects for inventories and other stock reported for the ( w^{th} ) day of the month (to denote the last day of the month enter 31). The option is not supported yet.</td>
</tr>
<tr>
<td></td>
<td>▪ Holidays – corresponds to the pre-defined trading day variables, modified to take into account country specific holidays. It means that after choosing this option the user should specify the type of the trading day effect (Calendar → tradingDays → TradingDays) and a previously defined calendar that includes the country specific holidays (Calendar → tradingDays → holidays).</td>
</tr>
<tr>
<td>Calendar → tradingDays → holidays</td>
<td>List of user-defined calendars to be used to create calendar regression variables. Option is available when Calendar → tradingDays → option is set to Holidays.</td>
</tr>
<tr>
<td></td>
<td>The user is expected to click the field to expand a list of previously defined calendars and choose an appropriate item.</td>
</tr>
<tr>
<td></td>
<td>The default setting is Default, which implies that the default calendar is used and no country-specific holidays are considered.</td>
</tr>
<tr>
<td>Calendar → tradingDays → userVariables</td>
<td>List of predefined regression variables to be included in the model. Option is available when Calendar → tradingDays → option is set to UserDefined.</td>
</tr>
<tr>
<td></td>
<td>When the user chooses the UserDefined type for the trading day effect estimation, one must specify the corresponding variables. It should be noted that such variables must have been previously defined (see 3.1.2), otherwise the list in the field Calendar → tradingDays → userVariables is empty.</td>
</tr>
<tr>
<td>Option</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
</tr>
<tr>
<td>Calendar → tradingDays → td&lt;br&gt;regression[variables=td]</td>
<td>Assigns a type of model-estimated regression effect to pre-specified regression variables. Option is available when Calendar → tradingDays → option is set to Default or Holidays. Acceptable values:&lt;br&gt;▪ None – means that no calendar variables will be included in the regression.&lt;br&gt;▪ TradingDays – includes six day-of-the-week regression variables.&lt;br&gt;▪ WorkingDays – includes the working/non-working day contrast variable.&lt;br&gt;The default setting is TradingDays.</td>
</tr>
<tr>
<td>Calendar → tradingDays → lp&lt;br&gt;regression[variables=lp]</td>
<td>Includes the leap-year effect in the model. Acceptable values:&lt;br&gt;▪ None – the leap-year effect is not included in the model;&lt;br&gt;▪ LeapYear – includes a contrast variable for the leap-year;&lt;br&gt;▪ LengthofPeriod – includes length-of-month (or length-of-quarter) as a regression variable to model the leap year effect.&lt;br&gt;The default setting is LeapYear.</td>
</tr>
<tr>
<td>Calendar → tradingDays → autoAdjust</td>
<td>Option for automatic correction for the leap year effect. It is available when in the Transformation section function is Auto. By default it is enabled.</td>
</tr>
<tr>
<td>Feature</td>
<td>Description</td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>Calendar → tradingDays → test</strong></td>
<td>Pre-tests the significance of the trading day regression variables using the AICC statistics.</td>
</tr>
<tr>
<td></td>
<td>The user can choose between the following options:</td>
</tr>
<tr>
<td></td>
<td>- Add – the trading day variables are not included in the initial regression model but they can be added to the RegARIMA model after the test.</td>
</tr>
<tr>
<td></td>
<td>- None – the test is not performed; the specified calendar variables are used in the model without pre-testing.</td>
</tr>
<tr>
<td></td>
<td>- Remove – the trading day variables belong to the initial regression model but they can be removed from the RegARIMA model after the test.</td>
</tr>
<tr>
<td></td>
<td>These options have no direct impact on the calendar tests themselves, but indirectly, through the definition of the regression model, on any previous test. For instance, on rare occasions the log/level test could be affected by add/remove on the trading day effect.</td>
</tr>
<tr>
<td></td>
<td>The default setting is Remove.</td>
</tr>
<tr>
<td><strong>Calendar → easter → IsEnabled</strong></td>
<td>The option enables the user to consider the Easter effect in the RegARIMA model. When the user ticks the option it means that the correction for the Easter effect is considered. The inclusion of the Easter effect to the model depends on the choice made in the Pretest section. Otherwise a correction for the Easter effect is not performed.</td>
</tr>
<tr>
<td><strong>Calendar → easter → Use Julian Easter</strong></td>
<td>When marked, it enables to include Easter, which date derives from Julian calendar and is converted to Gregorian calendar date. By default, the checkbox is unmarked.</td>
</tr>
<tr>
<td><strong>Calendar → easter → Pretest</strong></td>
<td>Pre-tests the significance of the Easter regression variable using the AICC statistics.</td>
</tr>
<tr>
<td></td>
<td>The user can choose between the following options:</td>
</tr>
<tr>
<td></td>
<td>- Add – the Easter effect is not included in the initial regression model but it can be added to the RegARIMA model after the test.</td>
</tr>
<tr>
<td></td>
<td>- None – the Easter effect is not pre-tested.</td>
</tr>
<tr>
<td></td>
<td>- Remove – the Easter effect belongs to the initial regression model and it can be removed from the RegARIMA model after the test.</td>
</tr>
<tr>
<td></td>
<td>These options have no direct impact on the calendar tests themselves, but indirectly, through the definition of the regression model, on any previous test. For instance, on rare occasions the log/level test could be affected by add/remove on the Easter effect.</td>
</tr>
<tr>
<td></td>
<td>The default setting is Add.</td>
</tr>
<tr>
<td>Option</td>
<td>Description</td>
</tr>
<tr>
<td>--------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Calendar → easter → Easter duration</td>
<td>Duration (length in days) of the Easter effect. The length of the Easter effect can range from 1 to 20 days. The Easter duration option is displayed when Calendar → easter → Pre-test is set to either None or Remove. The default value is 8.</td>
</tr>
</tbody>
</table>
| Pre-specified outliers                     | User-defined outliers are used when prior knowledge suggests that certain effects exist at known time points. Four pre-defined outlier types, which are simple forms of intervention variables, are implemented:  
  ▪ Additive Outlier (AO);  
  ▪ Level shift (LS);  
  ▪ Temporary change (TC);  
  ▪ Seasonal outliers (SO). Descriptions and formulas are available in 7.1.1. By default, no pre-specified outlier is included in the specification. |
| Intervention variables                     | The intervention variables are defined as in TRAMO. Following the definition, these effects are special events known a-priori (strikes, devaluations, political events, and so on). Intervention variables are modelled as any possible sequence of ones and zeros, on which some operators may be applied. Intervention variables are built as combinations of four basic structures:  
  ▪ Dummy variables;  
  ▪ Any possible sequence of ones and zeros;  
  \[
  \frac{1}{(1-\delta s B^s)^r} \frac{1}{(1-\delta B)} \left(0 < \delta \leq 1 \right), \\
  \frac{1}{(1-\delta B^s)} \left(0 < \delta \leq 1 \right), \\
  \frac{1}{(1-\delta)(1-B)} \left(0 < \delta \leq 1 \right), \\
  \frac{1}{(1-B)} \left(0 < \delta \leq 1 \right),
  \]
  \[B \text{ is backshift operator (i.e. } B^k X_t = X_{t-k}) \text{ and } s \text{ is frequency of the time series (} s = 12 \text{ for a monthly time series, } s = 4 \text{ for a quarterly time series).}
  
  \] These operations enable to generate not only AO, LS, TC, SO and RP outliers but also sophisticated intervention variables that are well-adjusted to the particular case. By default, no intervention variable is included in the specification. Intervention variables are not implemented in X-13ARIMA-SEATS, however they can be implemented in X-12-ARIMA program. |

---

30 In the original X-12-ARIMA program the length of the Easter effect can range from 1 to 25 days.
32 In the TRAMO/SEATS method this type of outlier is called transitory change.
34 Dummy variable is the variable that takes the values 0 or 1 to indicate the absence or presence of some effect.
be created by the user and introduced to the model as
userdefined variables.

<table>
<thead>
<tr>
<th>Ramps</th>
</tr>
</thead>
<tbody>
<tr>
<td>regression(variables = (rp))</td>
</tr>
<tr>
<td>A ramp effect means a linear increase or decrease in the level of the series over a specified time interval $t_0$ to $t_1$. All dates of the ramps must occur within the time series span. Ramps can overlap other ramps, additive outliers and level shifts. The graph and formula are available in 7.1.1. By default, no ramp is included in the specification.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The user-defined variable is an external regressor included by the user to the RegARIMA model. To add a user-defined variable to the model, one must specify the corresponding variable in the Variables window. First, click the Unnamed item and then click the Name field to expand a list of available variables and choose a variable from the list. It should be noted that such variables must have been previously defined (see 3.1.2), otherwise the list is empty. The user-defined regression variable associated to a specific component should not contain effects that have to be associated with another component.

Therefore, the following rules should be obeyed:

▪ The variable assigned to the trend or to the seasonally adjusted series should not contain a seasonal pattern;
▪ The variable assigned to the seasonal should contain neither a trend nor a level (i.e. should have a zero mean);
▪ The variable assigned to the irregular should contain neither a seasonal pattern nor a trend (i.e. should have a zero mean);

The user-defined variables should cover the period of the dependent series and the appropriate number of forecasts. For the other periods the variables are implicitly set to 0. If the forecasts are not provided, it will not alter the results of the seasonal adjustment but the forecasts of the final components will be unusable.

The effect of the user-defined variable can be assigned to the:

▪ Trend;
▪ Irregular component;
▪ Seasonal component;
▪ Seasonally adjusted series;
▪ Series;
▪ None of the above, which implies that the effect is an additional component (option Undefined), which is a default setting\textsuperscript{35}. With this option the regression variable is used to improve the modelling, but it is not removed from the series for the decomposition\textsuperscript{36}.

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
</table>

\textsuperscript{35} The convention for assigning effects for user-defined regressors is based on the TRAMO/SEATS solution. The original X-12ARIMA program does not provide the same functionality. The effect of additional component (option "Undefined") is a special feature introduced in the original X-12-ARIMA program and not available in TRAMO/SEATS.

\textsuperscript{36} A typical example could be a variable that measures the weather conditions: introducing them could improve the linearization; however, these effects are not allocated to a specific component. On the contrary, the impact of this variable will be split between all components.
The calendar component is not available in this section. Therefore, a user-defined variable assigned to the calendar effect should be added in the calendar part of the specification.

For the user-defined variable the structure of the lags can be specified using the options first lag and last lag. When the regression variable $x_t$ is introduced with first lag $l_a$ and last lag $l_b$, JDemetra+ includes in the TRAMO model a set of variables $x_{t-l_a}, \ldots, x_{t-l_b}$ and estimates the respective regression coefficients called the impulse response weights.

To include only the first lag ($x_{t-1}$) of the user-defined variable ($x_t$) the user should put first lag = last lag = 1. If for a monthly series one puts first lag = 0 and last lag = 11, it means that in addition to instantaneous effect of the user-defined variable, also the effects of 11 lagged explanatory variables are included in the model. In this case the set of estimated coefficients, called a transfer function, describe how the changes in $x_t$ that took place over a year are transferred to the dependent variable. However, the lagged variables are often collinear so that caution is needed in attributing much meaning to each coefficient.

By default, no user-defined variable is included in the specification.

### 6.1.14. Arima

An identification of the ARIMA part of the RegARIMA model can be done either in an automatic way or by the user who specifies the appropriate parameters. This choice is controlled by the **Automatic** option and results in a list of parameters dedicated for chosen ARIMA identification procedure. When the **Automatic** option is marked, the order of the ARIMA model results from the automatic identification procedure selects the final structure of the ARIMA model. The maximum orders of the regular polynomials are 3, and the maximum orders of seasonal polynomials are 1. The parameters available for automatic model identification are presented below.

#### Table 4.11: RegARIMA specification – options for the automatic identification of the ARIMA model.

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Automatic automdl</strong>;</td>
<td>When marked it enables automatic modelling of the ARIMA model to be performed.</td>
</tr>
</tbody>
</table>

---

37 More details and examples can be found in MARAVALL, A. (2008).
| **Accept Default** automdl; acceptdefault | Controls whether the default model (ARIMA(0,1,1)(0,1,1)) is chosen if the Ljung-Box Q-statistics for these model residuals is acceptable. If the default model is found to be acceptable, no further attempt will be made to identify a model or differencing order. By default, the **Accept Default** checkbox is unmarked. |
| **LjungBox limit** automdl; ljungboxlimit | Acceptance criterion for the confidence intervals of the Ljung-Box Q-statistic.³⁹ If the Ljung-Box Q-statistics for the residuals of a final model (checked at lag 24 if the series is monthly, 16 if the series is quarterly) is greater than **LjungBox limit**, the model is rejected, the outlier critical value is reduced, and model and outlier identification (if specified) is redone with a reduced value (see **Reduce CV** argument). The default parameter value is 0.95. |
| **ArmaLimit** automdl; armalimit | The threshold value for t-statistics of ARMA coefficients used for final test of model parsimony ⁴⁰. If the highest order ARMA coefficient has a t-value less than this value in magnitude, JDemetra+ will reduce the order of the model. The value given for **ArmaLimit** is also used for the final check for the significance of the constant term; if the constant term has a t-value less than **ArmaLimit** in magnitude, the program will remove the constant term from the set of regressors. The **ArmaLimit** value should be greater than zero. The default parameter value is 1. |
| **Balanced** automdl; balanced | Controls whether the automatic model procedure will have a preference for balanced models (i.e. models for which the order of the combined AR and differencing operator is equal to the order of the combined MA operator ⁴¹). By default, the **Balanced** checkbox is unmarked. When it is marked, the same preference as for the TRAMO program is used. |
| **Cancelation limit** automdl; – | Cancellation limit for the AR and the MA roots to be assumed equal ⁴². This option is mainly used in the automatic identification of the differencing order. The default parameter value is 0.1. |
| **Final UR (Diff.)** automdl; – | The threshold value for the final unit root test in the automatic identification of differencing order procedure. When one of the roots in the estimation of the (1,0,1)(1,0,1) plus mean model, which is performed in the second step of the automatic model identification procedure, is larger than **Final unit root limit**, in |

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Initial UR (Diff.)</strong></td>
<td>The threshold value for the initial unit root test in the automatic identification of differencing order procedure. When one of the roots in the estimation of the (2,0,0)(1,0,0) plus mean model, which is performed in the first step of the automatic model identification procedure, is larger than <strong>First unit root limit</strong>, in modulus, it is set equal to unity. The default parameter value is 1.0416666666666667.</td>
</tr>
<tr>
<td><strong>Reduce CV</strong></td>
<td>The percentage by which the outlier critical value will be reduced when the preferred model is found to have a Ljung-Box Q-statistic with an unacceptable confidence coefficient. The parameter should be between 0 and 1, and will only be active when automatic outlier identification is selected. The reduced critical value will be set to ((1-ReduceCV)\times CV), where CV is the original critical value. The default parameter value is 0.14268.</td>
</tr>
<tr>
<td><strong>Unit root limit</strong></td>
<td>The threshold value for the final unit root test. If the magnitude of an AR root for the final model is less than this number, a unit root is assumed, the order of the AR polynomial is reduced by one, and the appropriate order of the differencing (non-seasonal, seasonal) is increased. The parameter value should be greater than one. The default parameter value is 1.05.</td>
</tr>
</tbody>
</table>

When the **Automatic** checkbox in the **Arima** section is unmarked, JDemetra+ allows the user to specify the structure of the ARIMA part of the RegARIMA model. Initial values for the individual AR and MA parameters can be specified for the iterative estimation. Also, individual parameters can be held fixed at these initial values while the rest of the parameters are estimated. The options available here correspond to the original X-13ARIMA-SEATS `arima` spec with some limitations. JDemetra+ does not allow for operators with missing lags. Also the maximum lag is reduced in comparison with Win X-13.

JDemetra+ allows the user to fix individual parameters of the ARIMA model at initial values while the rest of the parameters are estimated. However, users should not specify initial values for the MA (moving average) parameters that yield the MA polynomial with the roots inside the unit

---
42 A unit root is an attribute of a statistical model of a time series whose autoregressive parameter is one. 46 See 7.6.1.3.
circle. It is also allowed, e.g. for the iterative estimation, to specify the initial values for the individual AR (autoregressive) and MA parameters.

Table 4.12: RegARIMA specification – options for the manual identification of the ARIMA model.

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Automatic</strong></td>
<td>When unmarked it enables the user to enter the parameters of the ARIMA model.</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>When marked it is considered that the mean is part of the ARIMA model (it highly depends on the chosen model).</td>
</tr>
<tr>
<td><strong>P</strong></td>
<td>The order of the non-seasonal autoregressive polynomial. The maximum order of the non-seasonal autoregressive polynomial is 4. The default value is 0.</td>
</tr>
</tbody>
</table>
| **phi**         | Coefficients of the non-seasonal autoregressive polynomial (AR). If used, to each non-seasonal AR parameter in the model a label that indicates the procedure of its estimation should be assigned. The choice can be made from:  
|                 | ▪ Undefined – estimates a parameter without the use of any user defined input (the default value).  
|                 | ▪ Initial – estimates a parameter using as initial condition the value defined by the user.  
|                 | ▪ Fixed – holds a parameter fixed during estimation at the value defined by the user.                                                               |
| **D**           | Non-seasonal differencing order. The maximum number of nonseasonal differences is 2. The default value is 1.                                     |
| **Q**           | The order of the non-seasonal moving average polynomial. The maximum order of the non-seasonal moving average polynomial is 4. The default value is 1. |
| **theta**       | Coefficients of the parameters of the non-seasonal, moving average polynomial (MA). If used, to each non-seasonal MA parameter in the model a label that indicates the procedure of its estimation should be assigned. The choice can be made from:  
|                 | ▪ Undefined – estimates a parameter without the use of any user defined input (the default value).  
|                 | ▪ Initial – estimates a parameter using as initial condition the value defined by the user.  
|                 | ▪ Fixed – holds a parameter fixed during estimation at the value defined by the user.                                                             |
### Option Description

<table>
<thead>
<tr>
<th><strong>BP arima; model</strong></th>
<th>The order of the seasonal autoregressive polynomial. The default value is 0.</th>
</tr>
</thead>
</table>
| **bphi** | Coefficients of the seasonal autoregressive polynomial (AR). If used, to each seasonal AR parameter in the model a label that indicates the procedure of its estimation should be assigned. The choice can be made from:
  - *Undefined* – estimates a parameter without the use of any user defined input (the default value).
  - *Initial* – estimates a parameter using as initial condition the value defined by the user.
  - *Fixed* – holds a parameter fixed during estimation at the value defined by the user. |
| **BD arima; model** | Seasonal differencing order. The maximum number of seasonal differences is 1. The maximum order of the seasonal autoregressive polynomial is 4. The default value is 1. |
| **BQ arima; model** | The order of the seasonal moving average polynomial. The maximum order of the seasonal moving average polynomial is 1. The default value is 1. |
| **btheta arima;** | Coefficients of the parameters of the seasonal moving average polynomial (MA). If used, to each seasonal MA parameter in the model a label that indicates the procedure of its estimation should be assigned. The choice can be made from:
  - *Undefined* – estimates a parameter without the use of any user defined input (the default value).
  - *Initial* – estimates a parameter using as initial condition the value defined by the user.
  - *Fixed* – holds a parameter fixed during estimation at the value defined by the user. |

### 6.1.1.5. Outliers

The *Outliers* section enables the user to perform an automatic detection of additive outliers, temporary change outliers, level shifts, seasonal outliers, or any combination of the four, using the specified model.

**Table 4.13: RegARIMA specification – options for the Outliers section.**

| Option | Description |
IsEnabled
outlier, –

Enables/disables the automatic detection of outliers in the span determined by the Detection span option. By default, the checkbox is marked which implies that the automatic identification of outliers is enabled.

Detection span→ type

A span of the time series to be searched for outliers. The available parameter’s values are:

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>outlier, span</td>
<td>▪ All – full time series span is considered in the modelling; ▪ From – date of the first time series observation included in the pre-processing model; ▪ To – date of the last time series observation included in the pre-processing model; ▪ Between – date of the first and the last time series observations included in the pre-processing model; ▪ Last – number of observations from the end of the time series included in the pre-processing model; ▪ First – number of observations from the beginning of the time series included in the pre-processing model; ▪ Excluding – number of observations excluded from the beginning (specified in the first field) and/or end (specified in the last field) of the time series in the pre-processing model. With the options Last, First, Excluding the span can be computed dynamically on the series. The default setting is All.</td>
</tr>
</tbody>
</table>

Use default critical value
outlier, critical

The critical value is automatically determined by the number of observations in the interval specified by the Detection span option. When Use default critical value is disabled, the procedure uses the critical value inputted in the Critical value item (see below). Otherwise, the default value is used (the first case corresponds to Critical value = "xxx"; the second corresponds to a specification without the critical argument). It should be noted that it is not possible to define a separate critical value for each outlier type.

By default, the checkbox is marked, which implies that the automatic determination of the critical value is enabled.

Critical value outlier; critical

The critical value used in the outliers’ detection procedure. The option is active once Use default critical value is disabled. By default, it is set to 4.0.

Additive outlier, ao

Automatic identification of additive outliers. By default, this option is enabled.
### Option Description

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level shift outlier, $ls$</td>
<td>Automatic identification of level shifts. By default, this option is enabled.</td>
</tr>
<tr>
<td>Transitory outlier, $tc$</td>
<td>Automatic identification of transitory changes. By default, this option is enabled.</td>
</tr>
<tr>
<td>Seasonal outlier, $ts$</td>
<td>Automatic identification of seasonal outliers. By default, this option is disabled.</td>
</tr>
<tr>
<td>TC rate outlier, $tcrate$</td>
<td>The rate of decay for the transitory change outlier. It must be a number greater than zero and less than one. The default value is 0.7.</td>
</tr>
<tr>
<td>Method outlier, method</td>
<td>Determines how the program successively adds detected outliers to the model. At present only one method, namely AddOne is supported. Using this approach JDemetra+ calculates t-statistics for each type of outlier specified (AO, TC and/or LS) at all time points for which outlier detection is being performed. If the maximum absolute outlier t-statistic exceeds the critical value, then an outlier has been detected and the appropriate regression variable is added to the model. The program then estimates the new model (the old model with the detected outlier added) and searches for an additional outlier. This process is repeated until no additional outliers are found. At this point, a backward deletion process is used to delete outliers, for which the absolute t-statistics no longer exceed the critical value, from the model. This is done one at a time beginning with the least significant outlier, until all outliers remaining in the model are significant. During backward deletion the usual (non-robust) residual variance estimate is used, which can yield somewhat different outlier t-statistics from those obtained during outlier detection.</td>
</tr>
<tr>
<td>LS Run outlier, $lsrun$</td>
<td>Computes a t-statistics to test the null hypotheses that all of the 2, 3, ..., $lsrun$ successive level shifts cancel to form a temporary level shift. If the value specified for $lsrun$ exceeds the total number of level shifts in the model following outlier detection, then $lsrun$ is reset to this total. The formula of temporary level shift is presented in 7.1.1. In the current version of JDemetra+ the option has no effect on the output. The default is 0, which indicates that no temporary level shift t-statistics are computed.</td>
</tr>
</tbody>
</table>

### 6.2. Documents

The **Documents** section, which belongs to the **Modelling** node of the **Workspace**, is designed to store results of modelling procedure performed with the TRAMO or the RegARIMA models. These
documents are displayed in the TramoDoc (for TRAMO) and the RegArimaDoc windows (for RegARIMA).

There are several ways to create the document. One of them is to choose from the main menu an option Statistical methods → Modelling → Single Analysis → Tramo/RegArima. By using this option a default specification will be assigned to this document (TRfull when Tramo is chosen, RG4 when RegArima is chosen). It means that the default specification will be used for modelling of the series inserted into document’s window.

Alternatively, the document can be created directly in the Workspace window, using the local menu options available for the selected specification. Expand the Modelling → specifications → tramo/regarima node and right click on the specification name to open the local menu and then chose an option Create document. An empty document will be added to the relevant place in the Modelling → documents section and the specification chosen by the user will be assigned to it.
An empty document can be also created from the Modelling → Documents level. It can be done by marking a modelling method in the navigation tree (tramo or regarima) and choosing an option New from the local menu.

![Figure 4.6: Creating a new document in the Documents section.](image)

All documents are added to the Modelling → Documents part of the Workspace window. Once the workspace is saved (see 3.1) all the documents defined for this workspace are saved as well. The user may then investigate the saved results of the modelling and update them using the Refresh data option from the TramoSeats Doc menu or the RegArima Doc menu, depending on the node, to which the specific document belongs to (tramo or regarima).

![Figure 4.7: Location of the documents for modelling with the RegARIMA model.](image)

To use the templates created directly in the Workspace window, double click on its name to display it. To perform an analysis, drag and drop the series from the Workspace window to the document window (for more details see 3.1.1).
Once the modelling is launched the output is generated automatically. The content of the document is generated automatically once the user drags and drops series into the relevant place of the document. The structure of each *tramo* and *regarima* document is the same. The results of the modelling are divided into two main parts: *Input* and *Model* and organised in a tree structure, which can be expanded. To investigate the specific node, click on it in the panel on the left to display its content in the panel on the right. Study the content of the section using the vertical scrollbar.
The fundamentals of the TRAMO and the RegARIMA models are the same. Therefore, the description of the output, which is presented in the subsequent sections, is illustrated with the results from the TRAMO model. However, it can be applied to the results from the RegARIMA model.

### 6.2.1. Input

*Input* includes the original series and the settings used to generate the output presented in the *Model* section.

![Figure 4.10: The view of the Specification panel.](image)

The original series are presented as a table in the *Series* node.
The table can be copied to Excel by dragging and dropping the top left corner cell to an Excel sheet. A standard local menu provides a list of available actions:

- **Open** – opens a window presenting time series as a graph and as a table.
- **Open with** – opens the time series in a separate window according to the user choice (chart & grid or only chart). The **All ts views** option is not currently available.
- **Rename** – enables the user to change the time series name.
- **Copy** – copies series and allows it to be pasted into another application e.g. into Excel.
- **Transpose** – changes the orientation of the table from horizontal to vertical.
- **Reverse chronology** – displays the series from last to first observation.
- **Single time series** – when it is marked observations are divided by calendar’s periods. Otherwise, data are presented as a standard time series.
- **Edit format** – allows the format used for displaying dates and values to be changed.
- **Use color scheme** – allows series values to be displayed in color.
- **Color scheme** – allows for a choice of color scheme from a pre-specified list.
- **Show bars** – presents a values in a table as horizontal bars.
- **Zoom** – an option for modifying the chart size.

**Paste, Select all and Clear** are disabled as they are not relevant for this view.

### 6.2.2. Model - generalities

The *Model* node includes basic information about the outcome of the model identification procedure and checking the goodness of fit. The summary information about the final model is available directly from the *Model* node. The content of this panel depends on the specifications used for processing and the results of the seasonal adjustment\(^\text{43}\).

---

\(^{43}\) For description of specifications see 4.1.
The first part contains fundamental information about the model.

**Summary**

*Estimation span:* [1-2000 - 2-2013]

158 observations

Series has been log-transformed

Trading days effects (6 variables)

No easter effect

3 detected outliers

Figure 4.13: The *Summary* section of the *Model* node.

*Estimation span* informs about the first and the last time series observation used for modelling. The notation of the estimation span varies according to the frequencies (for example, the span [2-1993 : 10-2006] represents a monthly time series and the span [II-1994 : I-2011] represents a quarterly time series). The message *Series has been log-transformed* is only displayed if a logarithmic transformation has been applied.

In the case of the pre-defined specifications: TR0, TR1, TR3, RG0, RG1 and RG3 no trading day effect is estimated. For TR2, RG2c, TR4 and RG4c pre-defined specifications, working day effects and the leap year effect are pre-tested and estimated if present. If the working day effect is significant, the pre-processing part includes the message *Working days effect* (1 regressor). The message *Working days effect* (2 regressors) means that the leap year effect has also been estimated. For TR5 and RG5c the trading day effect and the leap year effect are pre-tested. If the trading day effect has been detected, either of the messages *Trading days effect* (6 regressors) or *Trading days effect* (7 regressors) are displayed, depending whether the leap year effect has been detected or not. If the Easter effect is statistically significant, *Easter effect detected* is displayed.

In this section only the total number of detected outliers is visible. The additional information, i.e. type, location and coefficients of every outlier, can be found in the next section of the *Model* node.

The *Final Model* section informs about the outcome of the estimation process. *Number of effective observations* is the number of observations used to estimate the model, i.e. the number of observations of the transformed series (regularly and/or seasonally differenced) reduced by the *Number of estimated parameters*, which is the sum of regular and seasonal parameters for both autoregressive and moving average processes, mean effect, trading/working day effect, outliers, regressors and one.

*Likelihood* is a maximized value of Likelihood\(^44\) function after the iterations processed in Exact Maximum Likelihood Estimation, which is a method used to estimate the model. This value is used

---

\(^{44}\) The likelihood function is the joint probability (density) function of observable random variables. It is viewed as the function of the parameters given the realized random variables. Therefore, this function measures the probability of observing the particular set of dependent variable values that occur in the sample.
by model selection criteria: AIC, AICC, BIC (corrected by length) and Hannan-Quinn. Standard error of the regression (ML estimate) is the standard error of the regression from Maximum Likelihood Estimation.

An example of the output is presented in the chart below.

```
Final model

Likelihood statistics
Number of effective observations = 145
Number of estimated parameters = 12

Loglikelihood = 432.9975420418026
Transformation adjustment = -670.5332734005417
Adjusted loglikelihood = -237.5656313583912

Standard error of the regression (ML estimate) = 0.01212577242165636
AIC = 499.117126771767625
AICC = 501.55349980811146
BIC (corrected for length) = -8.447298960225401

Scores at the solution:
0.000379 -0.001998
```

Figure 4.14: The content of the Final model section.

Next the estimated values of model parameters (Coefficients), t-statistics (T-Stat) and corresponding p-values (P|T|>t) are displayed. JDemetra+ uses the following notation:

- **Phi(p)** – the \( p^{th} \) term in the non-seasonal autoregressive polynomial;
- **Theta(q)** – the \( q^{th} \) term in the non-seasonal moving average polynomial;
- **BPhi(P)** – the \( P^{th} \) term in the seasonal autoregressive polynomial;
- **BTheta(Q)** – the \( Q^{th} \) term in the seasonal moving average polynomial.

In the example below, the ARIMA model \((0,1,1)(0,1,1)\) was chosen, which means that one regular and one seasonal moving average parameter were calculated. The p-values indicate that **BTheta(1)** parameter is significant in contrast to the **Theta(1)** which is not significant.

---

45 AIC, AICC, BIC and Hannan-Quinn criteria are used by X-13ARIMA-SEATS while BIC (TRAMO definition) by TRAMO/SEATS. Information criteria formulas are given in 7.1.1.3. These criteria are used in seasonal adjustment procedures for the selection of the proper ARIMA model. The model with the smaller value of the model selection criteria is preferred. Maximum Likelihood Estimation is a statistical method for estimating the coefficients of a model. This method determines the parameters that maximize the probability (likelihood) of the sample data.

46 The variable is called statistically significant if it is so extreme that such a result would be expected to arise simply by chance only in rare circumstances (with the probability equal to p-value). Generally, the regressor is thought to be significant if p-value is lower than 5%.
For the fixed ARIMA parameters (see 4.1.1.4 or 4.1.2.4), JDemetra+ shows only the values of the parameters. Figure 4.16 presents the output from the manually chosen ARIMA model (2,0,0)(0,1,1) with fixed parameter \( B\theta_1 \). For the fixed parameter the \( T-Stat \) and \( (P|T|>t) \) are not displayed as no estimation is done for this parameter.

![Figure 4.15](image1)

### Arima model
\[(0.1.1)(0.1.1)\]

| Coefficients | T-Stat | \( P(|T|>t) \) |
|--------------|--------|----------------|
| \( \theta_1 \) | -0.106 | -1.16 0.2477 |
| \( \theta_2 \) | -0.4011 | -4.70 0.0000 |

Figure 4.15: The estimation’s results for ARIMA model.

If the ARIMA model contains a constant term (detected automatically or introduced by the user), the estimated value and related statistics are reported.

![Figure 4.16](image2)

### Arima model
\[(2.0.0)(0.1.1)\]

| Coefficients | T-Stat | \( P(|T|>t) \) |
|--------------|--------|----------------|
| \( \phi_1 \) | -1.2104 | -9.03 0.0000 |
| \( \phi_2 \) | 0.4517 | 3.37 0.0016 |
| \( \theta_2 \) | -0.5853 |                |

Figure 4.16: The results of the estimation of the ARIMA model with a fixed coefficient.

JDemetra+ presents estimated values of coefficients of one or six regressors depending on the type of calendar effect specification. In the case of working days effect one regressor is estimated.

![Figure 4.17](image3)

### Mean

| Coefficients | T-Stat | \( P(|T|>t) \) |
|--------------|--------|----------------|
| \( \mu \)   | 0.0980 | 15.16 0.0000 |

Figure 4.17: The results of the estimation of a mean effect.

When a trading days effect is estimated the Joint F-test value is reported under the table that presents estimated values. When the result of Joint F-test indicates that the trading day variables are jointly not significant the test result is displayed in red.

![Figure 4.18](image4)

### Working days

| Coefficients | T-Stat | \( P(|T|>t) \) |
|--------------|--------|----------------|
| \( \theta_3 \) | 0.0027 | 4.04 0.0001 |

Figure 4.18: The results of the estimation of a working day effect.
Figure 4.19: The results of the estimation of a trading day effect: the case of jointly not significant variables.

In the example below the RSA5c specification has been used and a trading day effect has been detected. In spite of the fact that some trading day regressors are not significant at the 5% significance level, the outcome of the joint F-test indicates that the trading day regressors are jointly significant (the F-test statistic is lower than 5%).

![Trading day results](image1)

**Joint F-Test = 1.62 (0.1474)**

If a leap year regressor has been used in the model specification, the value of the estimated leap year coefficient is also reported with the corresponding t-statistics and p-value. As the p-value presented on the picture below is greater than 0.05, it indicates that the leap year effect is not significant.

![Leap year results](image2)

**Joint F-Test = 4.01 (0.0003)**

If the option UserDefined (see 4.1.1.3 or 4.1.2.3) has been applied, JDemetra+ displays the User-defined calendar variables section with variables and their estimation results (the values of the parameters, corresponding t-statistics and p-values). The outcome of the joint F-test is displayed when more than one user-defined calendar variable is used.

![User-defined calendar variables](image3)
If the Easter effect is estimated, the following table will be displayed in the output. It is clear, that in the case presented below, Easter has a negative, significant effect on the time series.

![Easter Effect Table]

**JDemetra+** also presents the results of outlier detection. The table includes the type of outlier, its date, the value of the coefficient and corresponding t-statistics and p-values.

![Outlier Detection Table]

In all pre-defined specifications, except for TR0 and RG0, only additive outliers, temporary changes and level shifts are considered in the automatic outlier identification procedure. When seasonal outliers are also allowed, they appear in the same table as other outliers.

![Outlier Identification Procedure]

Results for pre-specified outliers (see 4.1.1.3 and 4.1.2.3) are displayed in a separate table.

![Pre-specified Outliers Table]
Regression variables are not identified automatically. They need to be defined by the user. The results of estimation of ramps that are pre-defined types of regression variables are displayed in a separate table. Available information concerning ramps, including the ramp spans, estimated coefficients and related statistics, is shown in the separate section.

| Coefficients | T-Stat | P(|t| > t) |
|--------------|--------|-----------|
| -0.0657      | -3.82  | 0.0002    |

Figure 4.27: The results of the estimation of the ramp effect.

All other intervention variables with corresponding statistics are shown under the Intervention variable(s) table.

| Coefficients | T-Stat | P(|t| > t) |
|--------------|--------|-----------|
| 1.5736       | 6.24   | 0.0000    |

Figure 4.28: The results of the estimation of the intervention variable.

User-defined variables are marked as Vars-1.x_1, Vars-1.x_2, ..., Vars-1.x_n and displayed in the separate tables.

| Coefficients | T-Stat | P(|t| > t) |
|--------------|--------|-----------|
| -0.0042      | -2.45  | 0.0157    |

| Coefficients | T-Stat | P(|t| > t) |
|--------------|--------|-----------|
| 0.0036       | 3.01   | 0.0031    |

Figure 4.29: The results of the estimation of the user-defined variables.

JDemetra+ also reports a list of missing observations, if any. JDemetra+ applies the AO approach to the estimation of the missing observations.

<table>
<thead>
<tr>
<th>Periods</th>
<th>Value</th>
<th>Standard error</th>
<th>Untransformed value</th>
</tr>
</thead>
<tbody>
<tr>
<td>9-1997</td>
<td>4.2967</td>
<td>0.0017</td>
<td>73.4568</td>
</tr>
<tr>
<td>1-1999</td>
<td>4.3796</td>
<td>0.0018</td>
<td>79.8059</td>
</tr>
</tbody>
</table>

Figure 4.30: The results of the estimation of the missing observations.

Detailed results are divided into several sections and are investigated in the sections 4.2.3 – 4.2.8 of this document.

---

47 Missing observations are treated as Additive Outliers and interpolated.
6.2.3. Forecasts

One of the important aspects of model checking is validation by forecasting out-of-sample values of the series\(^{48}\). Using the identified TRAMO model JDemetra+ produces point forecasts, forecast standard errors, and prediction interval. The prediction interval on the transformed scale is denoted as \textit{point forecast} \pm K \times \textit{forecast standard error}, where \(K\) is the standard error multiplier taken from a table of the normal distribution, corresponding to the specified coverage probability. JDemetra+ displays 95% prediction interval, which corresponds to \(K = 1.96\). The forecasting procedure assumes that no outliers appear in the forecasting period.

The graph presenting time series together with the values forecasted for the next year and corresponding prediction interval can be found directly in the \textit{Forecast} node.

![Figure 4.31: The Forecast panel with visual presentation of the estimated forecasts.](image)

The local menu offers the copy and export options, including sending the graph to the printer and save the graph as clipboard or a file in the PNG format. The \textit{Show precision gradient} option highlights the precision of the estimation using different shades of orange. As a rule, the precision decreases in time which is depicted by gradually more intense orange. The \textit{Copy all series} option enables the user to export time series together with the forecasts and the prediction intervals to another application.

---

Figure 4.32: The Forecast panel with visual presentation of the precision of the estimated forecasts.

The Out-of-sample test section, which is visible when the Forecast node is expanded, presents two tests. These tests check the performance of the out-of-sample forecasts, by comparison of forecast limits and the data. To perform this exercise, first the historical data sample is divided into a fit period and a test period (last 18 observations). For the series shortened by cutting off the last 18 observations the model is automatically selected and its parameters are fixed. The model is used to produce a one-period-ahead forecast (i.e. the forecast for $n + 1$, where $n$ is the length of the time series), and this estimation is performed 18 times for the gradually extended time series.

The first test compares forecast errors with the standard error of the in-sample residuals. The goodness of fit is assessed by checking if the mean of the forecast errors can be assumed zero. The second test compares mean squared forecast error with mean squared in-sample residuals. The result of the test is accepted when these two indicators are close to each other. The lack of consistency is clear evidence of model inadequacy.
6.2.4. Regressors

This section presents all deterministic regressors used by the TRAMO model, including the trading day variables, the leap year effect, outliers, the Easter effect, ramps, intervention variables and user-defined variables. To copy one series, make sure that the Single time series option from the local menu is unmarked (for local menu option see the description below Figure 4.11). Then click on its header and use the standard Ctrl+C shortcut. To copy whole table click on the top-left empty cell and use the standard Ctrl+C shortcut.
6.2.5. Arima

The Arima section demonstrates a theoretical spectrum\(^{49}\) of the stationary and non-stationary models. The local menu which is available for the graph, offers the copy and export options, including sending the graph to the printer and saving the graph as clipboard or as a file in the PNG format.

The Copy all visible option enables the user to export time series data to another application.

![Theoretical spectrum of the ARIMA model.](image)

In the bottom part the panel the ARIMA model used by TRAMO is presented using symbolic notation\(^{50}\) \((P, D, Q)(PB, DB, QB)\). Estimated coefficients of parameters (regular and seasonal AR and MA) are shown in closed form (i.e. using the backshift operator\(^{51}\) \(B\)). For each regular AR root (i.e. the solution of the characteristic equation) the argument and modulus are given. Details are given in see 7.1.2.1.

\[\text{Arima model (3,1,1)(0,1,1)}\]

\textbf{Polynomials}

- regular AR: 1,00000 - 0,284038 B - 0,0737544 B^2 - 0,175742 B^3
- regular MA: 1,00000 - 0,665024 B
- seasonal MA: 1,00000 - 0,655012 S

- \textbf{Regular AR inverse roots:}
  - argument=-2,0316, modulus=0,4951
  - argument=2,0316, modulus=0,4951
  - argument=0,0000, modulus=0,7228

![The details of ARIMA model used for modelling.](image)

---

\(^{49}\) Basic information about spectrum is presented in 7.3.

\(^{50}\) See 7.1.1.

\(^{51}\) Backshift operator \(B\) is defined as: \(B^k y_t = y_{t-k}\). It is used to denote lagged series.
For each regular AR root the argument and modulus are also reported (if present, i.e. if \( P > 0 \)) to inform to which time series component the regular roots would be assigned.

### 6.2.6. Pre-adjustment series

The table presented in this section contains series estimated by TRAMO. The contents of the table depend on the regressors included in the TRAMO model.

The following items can appear here:

- **Interpolated series** (yc) – series interpolated for the missing observations (if any);
- **Linearised series** (y_lin) – series adjusted for all deterministic effects, including a logarithmic transformation;
- **Series corrected for the calendar effect** (ycal) – series corrected for all calendar effects (also user-defined variables assigned to calendar component);
- **Deterministic component** (det) – all deterministic effects such as outliers, ramps, calendars etc.;
- **Calendar effect** (cal) – the total calendar effect, i.e. joint effect of moving holidays, trading/working day effect and the Easter effect;
- **Moving holidays effect** (omhe) – the same (provisionally) as Easter effect;
- **Trading day effect** (tde) – an automatically detected or user-entered trading day effect, i.e. predefined calendar effect, user-defined calendar and user-defined calendar regressors;
- **Easter effect** (ee) – an automatically detected or the user-defined Easter effect.

Figure 4.37: The example of the content of the Pre-adjustment panel.
• **Outliers effect on the trend component** (**out_t**) – level shifts effect; • **Outliers effect on the seasonal component** (**out_s**) – seasonal outliers effect; • **Outliers effect on the irregular component** (**out_i**) – additive and transitory change outliers effect;

• **Total outliers effect** (**out**) – the effects of the outliers on the trend, irregular and seasonal components;

• **Regression effect on the series** (**reg_y**) – user-defined variable effect assigned the series (the Component type option for the User-defined variables parameter is set to Series; see 4.1.1.3 or 4.1.2.3);

• **Regression effect on the seasonally adjusted series** (**reg_sa**) – the effect of user-defined variables effects assigned to the seasonally adjusted series (the Component type option for the User-defined variables parameter is set to Trend, Irregular and/or SeasonallyAdjusted; see 4.1.1.3 or 4.1.2.3);

• **Regression effect on the trend component** (**reg_t**) – the effect of ramps, intervention variables, for which $\Delta \neq 0$ and $\Delta S = 0$, and user-defined variables assigned to the trend component (the Component type option for the User-defined variables parameter is set to Trend; see 4.1.1.3 or 4.1.2.3);

• **Regression effect on the seasonal component** (**reg_s**) – the effect of intervention variables for which $\Delta S \neq 0$ and user-defined variables assigned to the seasonal component (the Component type option for the User-defined variables parameter is set to Series; see 4.1.1.3 or 4.1.2.3);

• **Regression effect on the irregular component** (**reg_i**) – the effect of user-defined variables effects assigned to irregular component (the Component type option for the User-defined variables parameter is set to Irregular; see 4.1.1.3 or 4.1.2.3);

• **Total regression effect** (**reg**) – the sum of the regression effects on the trend, seasonal component, irregular component, seasonally adjusted series and the separate regression effects assigned to none of components (in the last case the Component type option for the User-defined variables parameter should be set to Undefined; see 4.1.1.3 or 4.1.2.3).

### 6.2.7. Residuals

The examination of residuals from the fitted model for signs of model inadequacy is a crucial step of model validation. Therefore, JDemetra+ produces several residual diagnostics and graphs for detailed validation of the residuals. They are presented in the Residuals section and its subcategories.

---

52 If both $\Delta \neq 0$ and $\Delta S \neq 0$, an intervention variable is automatically assigned to the seasonal component.
The main panel in this section shows residuals from the TRAMO model in the graph and in the table. The graph can be used for visual examination of the residuals. If the model is adequate, the residuals should not contain any obvious, repetitive pattern.

The local menu which is available for the graph, offers the copy and export options, including sending the graph to the printer and saving the graph as clipboard or as a file in the PNG format. The Copy series option enables the user to select the series used to produce the graph and export it to another application.

![Graph showing residuals from the TRAMO model](image)

*Figure 4.38: Residuals from the TRAMO model presented in a graph.*

The lower part of the panel presents the values of the residuals.

![Table showing residuals from the TRAMO model](image)

*Figure 4.39: Residuals from the TRAMO model presented in a table.*

A standard local menu, which is available for this table, includes:

- **Select all** – copies series and allows it to be pasted into another application e.g. into Excel.
- **Transpose** – changes the orientation of the table from horizontal to vertical.
- **Reverse chronology** – displays the series from last to first observation.
▪ Single time series – when it is marked observations are divided by calendar’s periods. Otherwise, data are presented as a standard time series.
▪ Edit format – allows changing the format used for displaying dates and values.
▪ Use color scheme – allows series values to be displayed in colour.
▪ Color scheme – allows for a choice of colour scheme from a pre-specified list.
▪ Show bars – allows values to be displayed in a table as horizontal bars.
▪ Zoom – an option for modifying the chart size.

Paste and Clear are disabled as they are not relevant for this view.

6.2.7.1. Statistics

The aim of the residuals’ diagnostic is to check if it can be assumed that the residuals are random, normally distributed and do not contain any structures that can be successfully modelled. JDemetra+ produces various diagnostic statistics using the residuals from the TRAMO model to scrutinize their properties.

For each test the corresponding p-value is reported. A p-value is the probability of obtaining a test statistic at least as extreme as the one that was actually observed. The results are marked in green, yellow or red, depending on the result of statistical test used. Those in green denote that the problematic characteristic has not been detected (e.g. lack of normality of residuals, a significant autocorrelation in residuals). An outcome in yellow denotes the uncertain results. An outcome in red denotes cases where an issue should be addressed. Hence, test statistics will indicate the need to improve the model. Ideally, the model should be improved so that no test statistics indicate uncertainties in the results.

The normality of residuals is crucial for the validity of the prediction intervals produced in forecasting. To assess this property the Doornik-Hansen\(^ {53} \) test is applied (Normality). To give more insight into the outcome of this test also the closeness between the residuals mean, skewness and kurtosis is tested. A significant value of one of these statistics indicates that the standardized residuals do not follow a standard normal distribution. The results shown below reveal that the source of uncertain assessment of the normality is the poor skewness performance.

---

\(^ {53} \text{See 7.6.1.1.} \)
1. Normality of the residuals

<table>
<thead>
<tr>
<th></th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0,2216</td>
</tr>
<tr>
<td>Skewness</td>
<td>0,9940</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>0,1275</td>
</tr>
<tr>
<td>Normality</td>
<td>0,0182</td>
</tr>
</tbody>
</table>

4.40: The results of normality of residuals tests.

The independence is assessed by the results of the Ljung-Box Q-statistics and the Box-Pierce Q-statistics computed for regular and seasonal lags. These tests check for the presence of autocorrelation between lags, which is a sign that the values of residuals are not independent. The number of lags that are taken into account depends on the time series frequency. The test on regular lags examines the first 24 lags (for monthly series) or the first 16 lags (for quarterly series). The tests for residual seasonal autocorrelation in residuals consider first two seasonal lags, irrespective of the time series frequency.

Another test that checks for the presence of autocorrelation in the residuals is the Durbin-Watson statistic. A test outcome that is close to 2 indicates no sign of autocorrelation. For details of the Durbin–Watson test see 7.6.1.2.

2. Independence of the residuals

<table>
<thead>
<tr>
<th></th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ljung-Box(24)</td>
<td>0,1737</td>
</tr>
<tr>
<td>Box-Pierce(24)</td>
<td>0,2712</td>
</tr>
<tr>
<td>Ljung-Box on seasonality(2)</td>
<td>0,9786</td>
</tr>
<tr>
<td>Box-Pierce on seasonality(2)</td>
<td>0,9803</td>
</tr>
</tbody>
</table>

Durbin-Watson statistic: 2,1452

Figure 4.41: The results of independence of residuals tests.

The randomness in the signs of residuals is assessed by the Wald-Wolfowitz test, also called the Run test. It examines the hypothesis that a series of numbers is random. For data centred on the mean, the test calculates the number and length of runs. A run is defined as a set of sequential values that are either all above or all below the mean. An up run is a sequence of numbers each of which is above the mean; a down run is a sequence of numbers each of which is below the mean.

The test checks if the number of up and down runs are distributed equally in time. Both too many runs and too few runs are unlikely a real random sequence. The null hypothesis is that the values

---

54 See 7.6.1.3.
55 See 7.6.1.4.
56 See 7.6.1.5.
of the series have been independently drawn from the same distribution. The test also verifies the hypothesis that the length of runs is random.

### 3. Randomness of the residuals

<table>
<thead>
<tr>
<th>P-value</th>
<th>Runs around the mean: number</th>
<th>Runs around the mean: length</th>
<th>Up and Down runs: number</th>
<th>Up and Down runs: length</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0104</td>
<td></td>
<td>0.9906</td>
<td>0.9906</td>
<td>0.9906</td>
</tr>
</tbody>
</table>

Figure 4.42: The results of randomness of residuals tests.

The significant values of the Ljung-Box Q-statistics (or the Box-Pierce Q-statistics) of the squared residuals indicate the random variation of the coefficients or time-varying conditional variances. These effects cannot be modelled by TRAMO. Their presences cause the test statistics and forecast coverage intervals to have reduced reliability. The autocorrelation is checked by JDemetra+ for the first 24 lags (monthly series) or for the first 16 lags (quarterly series).

The example below shows the p-values marked in red, which indicate that the null hypothesis has been rejected. The outcome of the linearity of the residuals test provides evidence of autocorrelation in residuals which means that the linear structure is left in the residuals.

### 4. Linearity of the residuals

<table>
<thead>
<tr>
<th>P-value</th>
<th>Ljung-Box on squared residuals (24)</th>
<th>Box-Pierce on squared residuals (24)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0000</td>
<td></td>
<td>0.0001</td>
</tr>
</tbody>
</table>

Figure 4.43: The results of linearity of residuals tests.

The basic residuals characteristics are given in table 0-Statistics. They include sum of squares, mean squared error (MSE) and standard error of residuals.

#### 0 - Statistics

- **Sum of squares:** 0.0032
- **MSE:** 0.0002
- **Standard error:** 0.0134

Figure 4.44: Details of residuals basic characteristics.

Next section presents the detailed statistics and tests of the distribution of residuals.
To assess the independence in detail, the autocorrelation function\(^{57}\) of residuals with standard deviations, the Ljung-Box Q-statistics and Box-Pierce statistics are computed through each lag.

## 2. Independence tests

### Ljung-Box and Box-Pierce tests on residuals:

<table>
<thead>
<tr>
<th>Lag</th>
<th>Autocorrelation</th>
<th>Standard deviation</th>
<th>Ljung-Box test P-value</th>
<th>Box-Pierce test P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.2305</td>
<td>0.0774</td>
<td>0.0442</td>
<td>0.0488</td>
</tr>
<tr>
<td>2</td>
<td>0.1515</td>
<td>0.0774</td>
<td>0.0474</td>
<td>0.0515</td>
</tr>
<tr>
<td>3</td>
<td>0.0447</td>
<td>0.0774</td>
<td>0.0839</td>
<td>0.0783</td>
</tr>
<tr>
<td>4</td>
<td>0.0819</td>
<td>0.0774</td>
<td>0.1689</td>
<td>0.1739</td>
</tr>
<tr>
<td>5</td>
<td>0.1459</td>
<td>0.0774</td>
<td>0.1429</td>
<td>0.1739</td>
</tr>
<tr>
<td>6</td>
<td>0.0859</td>
<td>0.0774</td>
<td>0.0926</td>
<td>0.0631</td>
</tr>
<tr>
<td>7</td>
<td>0.1502</td>
<td>0.0774</td>
<td>0.0211</td>
<td>0.0273</td>
</tr>
<tr>
<td>8</td>
<td>0.0664</td>
<td>0.0774</td>
<td>0.0339</td>
<td>0.0463</td>
</tr>
<tr>
<td>9</td>
<td>0.2120</td>
<td>0.0774</td>
<td>0.0241</td>
<td>0.0333</td>
</tr>
<tr>
<td>10</td>
<td>0.0763</td>
<td>0.0774</td>
<td>0.0384</td>
<td>0.0523</td>
</tr>
<tr>
<td>11</td>
<td>0.0854</td>
<td>0.0774</td>
<td>0.0591</td>
<td>0.0643</td>
</tr>
<tr>
<td>12</td>
<td>0.0132</td>
<td>0.0774</td>
<td>0.0580</td>
<td>0.0783</td>
</tr>
<tr>
<td>13</td>
<td>0.1332</td>
<td>0.0774</td>
<td>0.0825</td>
<td>0.1097</td>
</tr>
<tr>
<td>14</td>
<td>0.0238</td>
<td>0.0774</td>
<td>0.1119</td>
<td>0.1451</td>
</tr>
<tr>
<td>15</td>
<td>0.0603</td>
<td>0.0774</td>
<td>0.1085</td>
<td>0.1437</td>
</tr>
<tr>
<td>16</td>
<td>0.1292</td>
<td>0.0774</td>
<td>0.1089</td>
<td>0.1473</td>
</tr>
<tr>
<td>17</td>
<td>0.0379</td>
<td>0.0774</td>
<td>0.1326</td>
<td>0.1869</td>
</tr>
<tr>
<td>18</td>
<td>0.0506</td>
<td>0.0774</td>
<td>0.1561</td>
<td>0.2050</td>
</tr>
<tr>
<td>19</td>
<td>0.0312</td>
<td>0.0774</td>
<td>0.1634</td>
<td>0.2464</td>
</tr>
<tr>
<td>20</td>
<td>0.0339</td>
<td>0.0774</td>
<td>0.2168</td>
<td>0.2993</td>
</tr>
<tr>
<td>21</td>
<td>0.0906</td>
<td>0.0774</td>
<td>0.1572</td>
<td>0.2733</td>
</tr>
<tr>
<td>22</td>
<td>-0.0261</td>
<td>0.0774</td>
<td>0.2367</td>
<td>0.3197</td>
</tr>
</tbody>
</table>

The presence of autocorrelation at seasonal lags is checked by the Ljung-Box Q-statistics and Box-Pierce Q-statistics and the respective p-values. They are accompanied by the autocorrelation function of residuals and its standard deviations.

\(^{57}\) See 7.6.

---

114
Ljung-Box and Box-Pierce tests on seasonal residuals:

<table>
<thead>
<tr>
<th>Lag</th>
<th>Autocorrelation</th>
<th>Standard deviation</th>
<th>Ljung-Box test</th>
<th>P-Value</th>
<th>Box-Pierce test</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>0.0972</td>
<td>0.1459</td>
<td>0.5086</td>
<td>0.4789</td>
<td>0.4441</td>
<td>0.4952</td>
</tr>
<tr>
<td>8</td>
<td>-0.2252</td>
<td>0.1459</td>
<td>0.5056</td>
<td>0.7764</td>
<td>0.4441</td>
<td>0.4952</td>
</tr>
</tbody>
</table>

Figure 4.47: The detailed results of independence of squared residuals tests.

The detailed results on various versions of the Run test are given in the next section.

3. Randomness

Runs around the mean

Number of values above the central line: 23
Number of values below the central line: 24

Runs: 25

<table>
<thead>
<tr>
<th>Test</th>
<th>Value</th>
<th>P-Value</th>
<th>Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>0.1507</td>
<td>0.3802</td>
<td>Normal(0,0;1,00)</td>
</tr>
<tr>
<td>Length</td>
<td>3.1205</td>
<td>1.0000</td>
<td>Chi2(47)</td>
</tr>
</tbody>
</table>

Up and down runs: 28

<table>
<thead>
<tr>
<th>Test</th>
<th>Value</th>
<th>P-Value</th>
<th>Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>-1.0585</td>
<td>0.2898</td>
<td>Normal(0,0;1,00)</td>
</tr>
<tr>
<td>Length</td>
<td>3.1032</td>
<td>1.0000</td>
<td>Chi2(46)</td>
</tr>
</tbody>
</table>

Figure 4.48: The detailed results of randomness of residuals tests.

Finally, the autocorrelation function of squared residuals with standard errors, the Ljung-Box Q-statistics and Box-Pierce Q-statistics computed through each lag can be inspected to assess the presence of residual autocorrelation structures.
4. Linearity tests

Ljung-Box and Box-Pierce tests on square residuals:

<table>
<thead>
<tr>
<th>Lag</th>
<th>Autocorrelation</th>
<th>Standard deviation</th>
<th>Ljung-Box test</th>
<th>P-Value</th>
<th>Box-Pierce test</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0312</td>
<td>0.1459</td>
<td>0.0494</td>
<td>0.8241</td>
<td>0.0484</td>
<td>0.9296</td>
</tr>
<tr>
<td>2</td>
<td>-0.0935</td>
<td>0.1459</td>
<td>0.0708</td>
<td>0.9052</td>
<td>0.0656</td>
<td>0.9977</td>
</tr>
<tr>
<td>3</td>
<td>-0.0202</td>
<td>0.1459</td>
<td>0.2129</td>
<td>0.5294</td>
<td>1.0454</td>
<td>0.5858</td>
</tr>
<tr>
<td>4</td>
<td>0.2000</td>
<td>0.1459</td>
<td>2.1444</td>
<td>0.6421</td>
<td>2.2036</td>
<td>0.6983</td>
</tr>
<tr>
<td>5</td>
<td>-0.0742</td>
<td>0.1459</td>
<td>3.2522</td>
<td>0.6510</td>
<td>2.2220</td>
<td>0.7274</td>
</tr>
<tr>
<td>6</td>
<td>-0.1147</td>
<td>0.1459</td>
<td>3.4368</td>
<td>0.7524</td>
<td>2.9719</td>
<td>0.8124</td>
</tr>
<tr>
<td>7</td>
<td>0.0565</td>
<td>0.1459</td>
<td>3.7403</td>
<td>0.8304</td>
<td>2.9682</td>
<td>0.8582</td>
</tr>
<tr>
<td>8</td>
<td>-0.0257</td>
<td>0.1459</td>
<td>5.1273</td>
<td>0.7439</td>
<td>4.2859</td>
<td>0.6357</td>
</tr>
<tr>
<td>9</td>
<td>-0.1654</td>
<td>0.1459</td>
<td>5.1699</td>
<td>0.8107</td>
<td>4.1210</td>
<td>0.8870</td>
</tr>
<tr>
<td>10</td>
<td>-0.06246</td>
<td>0.1459</td>
<td>5.8280</td>
<td>0.8285</td>
<td>4.7991</td>
<td>0.9042</td>
</tr>
<tr>
<td>11</td>
<td>0.1018</td>
<td>0.1459</td>
<td>7.0524</td>
<td>0.7324</td>
<td>5.8951</td>
<td>0.9299</td>
</tr>
<tr>
<td>12</td>
<td>0.1381</td>
<td>0.1459</td>
<td>7.3166</td>
<td>0.8359</td>
<td>5.8950</td>
<td>0.9230</td>
</tr>
<tr>
<td>13</td>
<td>-0.0591</td>
<td>0.1459</td>
<td>9.3003</td>
<td>0.7423</td>
<td>7.2537</td>
<td>0.8883</td>
</tr>
<tr>
<td>14</td>
<td>0.1726</td>
<td>0.1459</td>
<td>10.2943</td>
<td>0.7404</td>
<td>7.5348</td>
<td>0.8973</td>
</tr>
<tr>
<td>15</td>
<td>-0.1117</td>
<td>0.1459</td>
<td>12.4229</td>
<td>0.8468</td>
<td>0.1908</td>
<td>0.8673</td>
</tr>
</tbody>
</table>

Figure 4.49: The detailed results of linearity of residuals tests.

6.2.7.2. Distribution

The sample autocorrelation function (ACF) and sample partial autocorrelation function (PACF) of the (regularly and seasonally differenced, if necessary) time series $y_t$ are tools used in a well-established procedure of identification of an ARIMA model that originates from the Box-Jenkins method. The significant lags observed for those functions are a sign of autocorrelation or moving average processes present in the data. For identification of TRAMO models a modified approach is needed, as described in BELL, W.R., and HILLMER, S.C. (1984), since the presence of regression effects can distort the appearance of the ACF and PACF.

As model residuals are expected to be a random process, the ACF and PACF functions for residuals are expected not to contain significant values. The user is expected to examine these functions in the usual way to identify the AR and MA orders of the regression error term in the TRAMO model.

For a description of the autocorrelation function and the partial autocorrelation function, see 7.9.

The ACF and PACF graphs are accompanied by the picture that shows a residuals histogram and compares it with a normal distribution. It is expected that the residuals follow the normal distribution therefore the consecutive histogram values should be close to the relevant values of the normal distribution plot. The significant lags, however, indicate the presence of autocorrelation structure left in residuals.
6.2.8. Likelihood

The *Likelihood* panel gives some insights into the results of the estimation of the log-likelihood function. The chart represents the opposite of the log-likelihood function. The maximum log-likelihood is denoted by a red dot. The graph allows the user to study the parameters of the ARIMA model estimated in the optimization procedure. As the graph is three-dimensional, the graph is plotted for two chosen parameters. The user can modify this choice with the options available on the right.

The menu offers additional options for adjusting a view.
The settings for computation of likelihood function can be modified through the Parameters menu. The Steps parameter is the number of values used for each parameter. The Epsilon parameter defines the used range of the parameter. In other words, the likelihood function is computed, for each coefficient $c_i$ in the range $[c_i(\text{max}) - \epsilon; c_i(\text{max}) + \epsilon]$, for values separated by $(2\epsilon)(\text{steps} - 1)$. The value $c_i(\text{max})$ is the value of the coefficient corresponding to the max likelihood.

7. Seasonal adjustment

The primary aim of the seasonal adjustment process is to remove seasonal fluctuations from time series. To achieve this goal, seasonal adjustment methods decompose the original time series into components that capture specific movements. These components are: trend-cycle, seasonality and irregularity (see 7.1). The trend-cycle component includes long-term and medium-term movements in the data. For seasonal adjustment purposes there is no need to separate this component into two parts. JDemetra+ refers to the trend-cycle as trend and consequently this convention is used in this document.

This chapter presents the options of the seasonal adjustment processes performed by the methods implemented in JDemetra+ (X-12-ARIMA/X-13ARIMA-SEATS and TRAMO/SEATS) and discusses the output displayed by JDemetra+. As these seasonal adjustment methods use different approach to decomposition, the output produced for both of them have different structures and content. Therefore, the results for both methods are discussed separately. However, in contrast to the original programs, in JDemetra+ some quality indicators have been implemented for both methods, therefore their descriptions are not duplicated.
7.1. Specifications

The Seasonal adjustment section of the Workspace window contains a set of pre-defined specifications that enables the user to seasonally adjust the time series using two methods: TRAMO/SEATS and X-13ARIMA-SEATS. For both methods the parameters of the seasonal adjustment process can be set by the user. Specifications described in this section are sets of parameters and values assigned to them that contain all information necessary for seasonal adjustment. The default critical values used by the tests included in the specifications can be changed by the user in the Tools → Options menu (see 3.4.7).

The set of pre-defined specifications for seasonal adjustment encompasses the most commonly used sets of seasonal adjustment parameters. The names of these pre-defined specifications correspond to the terminology used in TSW+. The users are strongly recommended to start their analysis with one of those specifications (usually RSA4c or RSA5c for X-13ARIMA-SEATS and RSA4, RSA5 or RSAfull for TRAMO/SEATS) then, if need be, to change some of the options afterwards using the Specification button (see the ‘JDemetra+ User Guide’ (2016), 3.2.1). The default specification for TRAMO/SEATS and multi-documents is RSAfull, while for X-13ARIMA-SEATS it is RSA4c.

Table 7.1: Pre-defined seasonal adjustment specifications.

<table>
<thead>
<tr>
<th>Specification</th>
<th>Transformation</th>
<th>Pre-adjustment for leap-year</th>
<th>Working days</th>
<th>Trading days</th>
<th>Easter effect</th>
<th>Outliers</th>
<th>ARIMA model</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSA0</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>(0,1,1)(0,1,1)</td>
</tr>
<tr>
<td>RSA1</td>
<td>test</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>test</td>
<td>(0,1,1)(0,1,1)</td>
</tr>
<tr>
<td>RSA2</td>
<td>test</td>
<td>no</td>
<td>test</td>
<td>no</td>
<td>test</td>
<td>test</td>
<td>(0,1,1)(0,1,1)</td>
</tr>
<tr>
<td>RSA3</td>
<td>test</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>test</td>
<td>AMI</td>
</tr>
<tr>
<td>RSA4</td>
<td>test</td>
<td>no</td>
<td>test</td>
<td>no</td>
<td>test</td>
<td>test</td>
<td>AMI</td>
</tr>
<tr>
<td>RSA5</td>
<td>test</td>
<td>no</td>
<td>no</td>
<td>test</td>
<td>test</td>
<td>test</td>
<td>AMI</td>
</tr>
<tr>
<td>RSAfull</td>
<td>test</td>
<td>no</td>
<td>test</td>
<td>test</td>
<td>test</td>
<td>test</td>
<td>AMI</td>
</tr>
<tr>
<td>X11</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>(0,1,1)(0,1,1)</td>
</tr>
<tr>
<td>RSA1</td>
<td>test</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>test</td>
<td>(0,1,1)(0,1,1)</td>
</tr>
<tr>
<td>RSA2c</td>
<td>test</td>
<td>test</td>
<td>test</td>
<td>no</td>
<td>test</td>
<td>test</td>
<td>(0,1,1)(0,1,1)</td>
</tr>
<tr>
<td>RSA3</td>
<td>test</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>test</td>
<td>AMI</td>
</tr>
</tbody>
</table>
Explanations for settings:

- **Transformation test** – a test is performed to choose between additive decomposition (no transformation) and multiplicative decomposition (logarithmic transformation).

- **Pre-adjustment for leap-year** – a correction of the February values applied to the original series before logarithmic transformation. The original values in February are multiplied by \( \frac{28.25}{29} \) for leap years and by \( \frac{28.25}{28} \) for non-leap years. Values for other months are not modified.

- **Working days** – a pre-test is made for the presence of the working day effect by using one parameter specification.

- **Trading days** – a pre-test is made for the presence of the trading day effect by using six parameters specification.

- **Easter** – the tests for the necessity of a correction for the Easter effect in the original series. The length of the Easter effect is 6 days (for TRAMO/SEATS specifications) and 8 days (for X-13ARIMA-SEATS specifications).

- **Outliers** – an automatic identification of three types of outlier: AO (additive outliers), LS (level shifts), TC (transitory changes) using a default critical value.

- **ARIMA model** – the choice between fixing the ARIMA model structure to \((0,1,1)(0,1,1)\) or searching for the ARIMA model using automatic model identification procedure (AMI). The \((0,1,1)(0,1,1)\) model (called the Airline model) is used as a default model in several TRAMO/SEATS and X-13ARIMA-SEATS specifications because it has been shown in many studies that this model is appropriate for many real seasonal monthly or a quarterly time series. Moreover, the Airline model approximates well many other models and provides an excellent "benchmark" model\(^8\).

The user may add new seasonal adjustment specifications to the Workspace window. To do it, go to the Seasonal adjustment section, right click on the tramoseats or x13 item in the specifications node and select New from the local menu.

---

\(^8\) MARAVALL, A. (2009).
7.1.1. TramoSeats

This section discusses the options available for the TRAMO/SEATS specifications, which are based on the original program developed by Agustin Maravall and Victor Gómez. It is divided into five parts that correspond to the TRAMO/SEATS specification sections and are presented in the order in which they are displayed in the graphical interface of JDemetra+.

To avoid unnecessary repetitions, the description of Series, Estimate, Transformation, Regression, Outliers and Arima nodes is omitted, as it is provided in the 4.1.1. Therefore, this section focuses on the decomposition part of the seasonal adjustment process and the benchmarking options.

To facilitate the comparison between JDemetra+ specifications and specifications used in TSW+, under each option the name of the corresponding parameter from the original software is given, if any. For an exact description of the different parameters, the user should refer to the documentation of the original TRAMO/SEATS program. For each parameter the default parameter value, which is displayed for a template created in the Workspace window is reported (in the Workspace window go to the Seasonal Adjustment section, right click on the tramoseats item in the specifications node and select New from the local menu).
For the pre-defined specifications the items are fixed, while in the case of the user-defined specification the user can set them individually. However, as in some cases the choice of a given value results in limitation of the possible alternatives for other parameters, the user is not entirely free to set the parameters values.

7.1.1.1. Series

In the context of seasonal adjustment it is usually assumed that long time series are those exceeding twenty years of length. Performing seasonal adjustment of long time series can be difficult. Over such a long period the underlying data generating process can change determining changes also in the components and in the components structure. In this case to perform the adjustment over the whole series can produce sub-optimal results mainly in the most recent and the initial parts of the series. Therefore it is reasonable to limit long time series to the most recent observations\(^5\). The Series section allows the user to limit the span (data interval) of the data to modelled or seasonally adjusted.

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Series span → type</strong></td>
<td>Specifies the span (data interval) of the time series to be used for the seasonal adjustment process. When the user limits the original time series to a given span only this span will be used in the computations. The available parameters for this option are:</td>
</tr>
<tr>
<td></td>
<td>- <em>All</em> – full time series span is considered in the modelling;</td>
</tr>
<tr>
<td></td>
<td>- <em>From</em> – date of the first time series observation included in the pre-processing model;</td>
</tr>
<tr>
<td></td>
<td>- <em>To</em> – date of the last time series observation included in the pre-processing model;</td>
</tr>
<tr>
<td></td>
<td>- <em>Between</em> – date of the first and the last time series observations included in the pre-processing model;</td>
</tr>
<tr>
<td></td>
<td>- <em>Last</em> – number of observations from the end of the time series included in the pre-processing model;</td>
</tr>
<tr>
<td></td>
<td>- <em>First</em> – number of observations from the beginning of the time series included in the pre-processing model;</td>
</tr>
<tr>
<td></td>
<td>- <em>Excluding</em> – number of observations excluded from the beginning (specified in the <em>first</em> field) and/or end (specified in the <em>last</em> field) of the time series in the pre-processing model.</td>
</tr>
</tbody>
</table>

With the options *Last*, *First*, *Excluding* the span can be computed dynamically on the series. The default setting is *All*.

Series span → Preliminary Check

When marked, it checks the quality of the input series and excludes from processing highly problematic ones: e.g. those with a number of outliers, identical observations and/or missing values above the respective threshold values. When unmarked, the thresholds are ignored and process is performed, when possible.

By default, the checkbox is marked.

7.1.1.2. Seats

This section includes the settings relevant for the decomposition step, performed by the SEATS algorithm. A basic description of the SEATS method can be found in 7.1.2.

Table 7.2: TRAMO/SEATS specification – options for the Seats section.

<table>
<thead>
<tr>
<th>Item</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approximation mode</td>
<td>In general, SEATS decomposes the ARIMA model received from TRAMO. On some occasions, the ARIMA model identified by TRAMO results in so called a non-admissible decomposition i.e. a decomposition for which the condition that all components have a non-negative spectrum for all frequencies has failed. In such cases an approximation might be use to choose an acceptable ARIMA model. The available actions that can be performed in the case of a non-admissible decomposition are:</td>
</tr>
</tbody>
</table>
| Seats parameters; noadmiss | - *None* – when the model does not accept an admissible decomposition, no approximation is made which means that no decomposition is performed by SEATS;  
- *Legacy* – when the model does not accept an admissible decomposition, it is automatically replaced with a decomposable one. The forecasts of the components obtained in SEATS with the new ARIMA model (sum of the components forecasts) will not add to the series forecast of the model passed by TRAMO.  
- *Noisy* – the new ARIMA model is obtained by adding white noise to the non-admissible model estimated by TRAMO. In this case, the forecasts of the series in TRAMO and in SEATS are the same; also the sum of the components forecasts is the same as the forecast of the series with the TRAMO model.  
*The default setting is* Legacy.  |

60 See 7.1.2.
| **MA unit root boundary**  
*Seats parameters: xl* | A parameter to control the AR and MA roots of the model. When the modulus of a root converges within an interval around 1, the program automatically fixes the root. More specifically, when the modulus of an estimated root falls in the range \((xl, 1)\), it is set to 1 if it is a root in the AR polynomial. If a root is in the MA polynomial, it is set to \(xl\). The default value is 0.95. |
|---|---|
| **Trend boundary**  
*Seats parameters: rmod* | The trend boundary is defined for the modulus of the inverse of the real AR roots. If the modulus of the inverse of the real root is greater than the **Trend boundary**, the AR root is integrated into the trend component. Otherwise the root is integrated into the seasonal component or transitory component (see **Seasonal boundary**). The default parameter value is 0.5. |
| **Seasonal boundary**  
*Seats parameters: smod* | The seasonal boundary is defined for the modulus of the inverse of the real negative AR roots. If the modulus of the inverse negative real root is greater (or equal) than **Seasonal boundary**, the AR root is integrated into the seasonal component. Otherwise the root is integrated into the trend or transitory component (see **Trend boundary**). The default parameter value is 0.5. |
| **Seasonal tolerance**  
*Seats parameters: epsphi* | The tolerance (measured in degrees) to allocate the AR non-real roots to the seasonal component (if the modulus of the inverse complex AR root is greater than the **Trend boundary** and the frequency of this root differs from one of the seasonal frequencies by less than **Seasonal tolerance**) or the transitory component (otherwise). The default parameter value is \(\frac{\pi}{90}\) rad (2 degrees). |
The estimation method of the unobserved components.

Options:

- **Burman** – the algorithm which is used by the original TRAMO/SEATS method and is the default option for JDemetra+. Although it is the most efficient one, it cannot handle MA unit roots and it may become numerically unstable when some roots of the MA polynomial are near 1. In such cases the Wiener-Kolmogorov approach may lead to a significant underestimation of the standard deviations of the components.

- **KalmanSmoother** – the most robust algorithm. It is not disturbed by (quasi-) unit roots in MA. It is slightly slower than the Burman’s algorithm. It should also be noted that it provides exact measures of the standard errors of the estimates (identical to the McElroy’s results).

- **McElroyMatrix** – the algorithm, which is much slower than other options and presents the same stability issues as the Burman’s algorithm. However, it provides additional results (full covariance matrix of the estimates) that may be useful.

<table>
<thead>
<tr>
<th>Item</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is enabled</td>
<td>Enables the user to perform benchmarking. By default, the checkbox is unmarked.</td>
</tr>
<tr>
<td>Target</td>
<td>Specifies the target variable for the benchmarking procedure. Options:</td>
</tr>
<tr>
<td></td>
<td>- <strong>Original</strong> – the raw time series are considered as target data;</td>
</tr>
<tr>
<td></td>
<td>- <strong>Calendar Adjusted</strong> – the time series adjusted for calendar effects are considered as target data.</td>
</tr>
</tbody>
</table>

The default setting is **Original**.

---

61 The disturbance filter of Koopman is nearly as fast as the Burman’s solution. However, it does not provide the standard deviations of the estimates.
Use forecast

The forecasts of the seasonally adjusted series and of the target variable (Target) are used in the benchmarking computation so the benchmarking constraint is applied also to the forecasting period. By default, the checkbox in unmarked (forecasts are not used).

Rho

The value of the AR(1) parameter (set between 0 and 1). The default value of 1 is equivalent to Denton benchmarking.

Lambda

A parameter that relates to the weights in the regression equation; it is typically equal to 0, 1/2 or 1. A parameter equal to 1 (default value) makes the method equivalent to multiplicative benchmarking, while a parameter equal to 0 makes the method equivalent to additive benchmarking.

7.1.2. X13

This section discusses the options available for the X-13ARIMA-SEATS specifications, which are based on the original X-13ARIMA-SEATS program developed by U.S. Census Bureau. The X13ARIMA-SEATS specifications are - to a very large extent - organised according to the different individual specifications of the original program and are presented in the order in which they are displayed in the graphical interface of JDemetra+.

Figure 7.3: A list of the X-13ARIMA-SEATS specification’s sections.

To avoid unnecessary repetitions, the description of Series, Estimate, Transformation, Regression, Arima and Outliers nodes is omitted, as it is provided in section 4.1.1. Therefore, this section focuses on the decomposition part of the seasonal adjustment process and the benchmarking options.

To facilitate the comparison between JDemetra+ specifications and specifications used in Win X13, under each option the name of the corresponding specification and argument from the original software is given, if any. For an exact description of the different parameters, the user should refer
to the documentation of the original X-13ARIMA-SEATS program. For each parameter the default parameter value, which is displayed for a template created in the Workspace window is reported (in the Workspace window go to the Seasonal Adjustment section, right click on the x13 item in the specifications node and select New from the local menu).

For the pre-defined specifications the items are fixed, while in the case of the user-defined specification the user can set them individually. However, as in some cases the choice of a given value results in limitation of the possible alternatives for other parameters, the user is not entirely free to set the parameters values.

7.1.2.1. Series

In the context of seasonal adjustment it is usually assumed that long time series are those exceeding twenty years of length. Performing seasonal adjustment of long time series can be difficult. Over such a long period the underlying data generating process can change determining changes also in the components and in the components structure. In this case to perform the adjustment over the whole series can produce sub-optimal results mainly in the most recent and the initial parts of the series. Therefore it is reasonable to limit long time series to the most recent observations. The Series section allows the user to limit the span (data interval) of the data to modelled or seasonally adjusted.

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
</table>

---

Series span → type

Specifies the span (data interval) of the time series to be used for the seasonal adjustment process. When the user limit the original time series to a given span only this span will be used in the computations. The available parameters for this option are:

- **All** – full time series span is considered in the modelling;
- **From** – date of the first time series observation included in the pre-processing model;
- **To** – date of the last time series observation included in the pre-processing model;
- **Between** – date of the first and the last time series observations included in the pre-processing model;
- **Last** – number of observations from the end of the time series included in the pre-processing model;
- **First** – number of observations from the beginning of the time series included in the pre-processing model;
- **Excluding** – number of observations excluded from the beginning (specified in the first field) and/or end (specified in the last field) of the time series in the pre-processing model.

With the options Last, First, Excluding the span can be computed dynamically on the series. The default setting is All.

Series span → Preliminary Check

When marked, it checks the quality of the input series and excludes from processing highly problematic ones: e.g. these with a number of outliers, identical observations and/or missing values above the respective threshold values. When unmarked, the thresholds are ignored and process is performed, when possible. By default, the checkbox is marked.

7.1.2.2. X11

This section includes the settings relevant for the decomposition step, performed by the X11 algorithm. A basic description of the X11 method can be found in 7.1.3.

Table 7.4: X13 specification – options for the X11 section.

<table>
<thead>
<tr>
<th>Item</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mode</strong></td>
<td>Time series can be regard as a composition of the distinctive components: trend-cycle, seasonal component, calendar component and irregular movements. For the decomposition purpose some assumption need to be made concerning the aggregation</td>
</tr>
<tr>
<td>Item</td>
<td>Comments</td>
</tr>
<tr>
<td>------</td>
<td>----------</td>
</tr>
</tbody>
</table>
| function that combines the components to form the linearized (stochastic) time series. The **Mode** parameter determines the mode of the seasonal adjustment decomposition to be performed. The choice can be made between:  
  - *Undefined* – no assumption concerning the relationship between the time series components is made.  
  - *Additive* – decompose an original time series into a sum of the components.  
  - *Multiplicative* – decompose an original time series assuming a multiplicative relationship between the components. It requires a bias correction for its trend and SA estimates.  
  - *LogAdditive* – performs an additive decomposition of the logarithms of the series being adjusted. It requires a bias correction for its trend estimates (due to geometric means being less than arithmetic means) as well as a different calibration for extreme value identification based on the log normal distribution.  

If **Transformation** is set to **Log** (see 4.2.3) the **Mode** parameter should be set to *Undefined* (*Multiplicative* or *LogAdditive* are allowed, too). If the transformation is set to **None** the **Mode** parameter should be set to *Additive*. If the transformation is set to **Auto**, the **Mode** parameter is automatically set to *Undefined*. The setting chosen by the user may be changed by the program automatically, if needed. The default setting is *Undefined*. |
| Seasonal component | When the checkbox is marked JDematra+ computes a seasonal component. Otherwise, the seasonal component is not estimated and its values are set it to 0 (additive decomposition) or 1 (multiplicative decomposition). By default, the checkbox is marked. |
| **Forecast horizon** | **forecast; maxlead** | Length of the forecasts generated by the RegARIMA model in months (positive values) or years (negative values). If **Forecast horizon** is set to 0 the X-11 procedure does not use any modelbased forecasts but the original X-11 type forecasts for one year. |

---

63 When the series is logarithmically transformed, the annual mean of original series is greater than the mean of seasonally adjusted series (and of trend). It is due to the fact that the geometric mean is less than arithmetic mean. For this reason, if the bias is large the correction should be applied to avoid such discrepancies.

LSigma
\textit{x11; sigmalim}

Specifies the lower sigma limit used to down weight extreme irregular values in the internal seasonal adjustment iterations. Valid list values are any real numbers greater than zero with the lower sigma limit less than the upper sigma limit. The default value is 1.5.

USigma
\textit{x11; sigmalim}

Specifies the upper sigma limit used to down weight extreme irregular values in the internal seasonal adjustment iterations.

<table>
<thead>
<tr>
<th>Item</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Valid list values are any real numbers greater than zero with the lower sigma limit less than the upper sigma limit. The default value is 2.5.</td>
</tr>
</tbody>
</table>

Seasonal filter \textit{x11; seasonalma}

Specifies which seasonal moving average\textsuperscript{65} (i.e. seasonal filter) will be used to estimate the seasonal factors for the entire series. The following filters are available\textsuperscript{66}:

- $S_3 \times 1$ – $3 \times 1$ moving average.
- $S_3 \times 3$ – $3 \times 3$ moving average.
- $S_3 \times 5$ – $3 \times 5$ moving average. $S_3 \times 9$ – $3 \times 9$ moving average.
- $S_3 \times 15$ – $3 \times 15$ moving average.
- Stable – a single seasonal factor for each calendar period is generated by calculating the simple average of all the values for each period (taken after detrending and outlier correction).
- X11\textit{Default} – $3 \times 3$ moving average is used to calculate the initial seasonal factors and a $3 \times 5$ moving average to calculate the final seasonal factor.
- Msr – automatic choice of seasonal filter. The seasonal filters can be selected for the entire series, or for a particular month or quarter.

The default value is Msr.

\textsuperscript{65} See 7.1.4.1 for a description of moving averages.
### Details on seasonal filters

*Period specific seasonal filters are offered as an option in X-11 in order to account for seasonal heteroskedasticity (see 7.1.3). This option enables the user to assign different seasonal filters to each period. This option is enabled only after executing a seasonal adjustment process with settings described in the specification because only then the frequency of the series, which is necessary to define the filters, is known. By default, this item is empty.*

### Automatic Henderson filter

*Automatic selection of the length of the Henderson filter is performed when the corresponding item is selected. Otherwise, the length given by the user in the **Henderson filter** item is used. By default, the checkbox is marked. The length of Automatic Henderson filter is 13.*

### Henderson filter

*Enables the user to apply the user-defined length of the Henderson filter. The option is available when the **Automatic Henderson filter** checkbox is unmarked. The default setting is 13. The length of the filters can be set as any odd number between 3 and 101.*

### Calendarsigma

*Specifies if the standard errors used for extreme values detection and adjustment are computed separately for each calendar*

<table>
<thead>
<tr>
<th>Item</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Item</strong></td>
<td><strong>Comments</strong></td>
</tr>
<tr>
<td>Calendarsigma</td>
<td>month/quarter (<em>All</em>); or separately for two complementary sets of calendar months/quarters specified by the sigmavec parameter (<em>Select</em>). Other options are to compute the standard errors separately for each period only if Cochran’s hypothesis test determines that the irregular component is heteroskedastic by calendar month/quarter (<em>Signif</em>) or to compute them from 5 year spans of irregulars (<em>None</em>). The default value is <em>None.</em></td>
</tr>
<tr>
<td>Sigmavec</td>
<td>The parameter is displayed only if <strong>Calendarsigma</strong> is set to <em>Select</em>. It specifies one of the two groups of periods (months or quarters) for whose irregulars a group standard error will be calculated. For each period the user sets the parameter value either to <em>Group1</em> or to <em>Group2</em>. By default, for all periods the parameter is set to <em>Group1.</em></td>
</tr>
</tbody>
</table>

---

67 See 5.2.2.2.3.
When the checkbox is marked, forecasts and backcasts from the RegARIMA model are not used in the generation of extreme values in the seasonal adjustment routines. Otherwise, the full forecast and backcast extended series are used in the extreme value process. By default, the checkbox is unmarked.

7.1.2.3. Benchmarking

The Benchmarking section allows the forcing of the annual sums of the seasonally adjusted data to be equal to the annual sums of the raw or calendar adjusted data. For description of benchmarking see 7.8.

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is enabled</td>
<td>Enables the user to perform benchmarking. By default, the checkbox is unmarked.</td>
</tr>
<tr>
<td>Target</td>
<td>Specifies the target variable for the benchmarking procedure. Options:</td>
</tr>
<tr>
<td></td>
<td>▪ Original – the raw time series are considered as target data;</td>
</tr>
<tr>
<td></td>
<td>▪ Calendar Adjusted – the time series adjusted for calendar effects are considered as target data. The default setting is Original.</td>
</tr>
<tr>
<td>Use forecast</td>
<td>The forecasts of the seasonally adjusted series and of the target variable (Target) are used in the benchmarking computation so the benchmarking constrain is applied also to the forecasting period. By default, the checkbox in unmarked.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rho</td>
<td>The value of the AR(1) parameter (set between 0 and 1). By default it is equal to 1.</td>
</tr>
<tr>
<td>Lambda</td>
<td>A parameter that relates to the weights in the regression equation; it is typically equal to 0, 1/2 or 1. By default it is equal to 1.</td>
</tr>
</tbody>
</table>

7.2. Documents

The Documents section, which belongs to the Seasonal Adjustment node, is designed to store results of the seasonal adjustment process resulting from the TRAMO/SEATS and X-13ARIMA-SEATS methods. These documents are displayed in the TramoSeatsDoc windows (for a time series adjusted with TRAMO/SEATS), the RegArimaSeatsDoc window (for a time series adjusted with X-13ARIMASEATS) and the SAProcessing window (for a set of series adjusted with one of both methods).
There are several ways to create this document. One of them is to choose the relevant option from the main menu. For an analysis of a single time series select *Statistical methods* → *Seasonal Adjustment* → *Seasonal adjustment* → *Single analysis* → *TramoSeats/X13* and follow the ‘*JDemetra+ User Guide*’ (2016), case study 3.1.1.

This option implies that a default specification will be assigned to the document (*TRfull* when *TramoSeats* is chosen, *RSA5c* when *X13* is chosen). It means that a default specification will be used for seasonal adjustment of the series inserted into document’s window.

For an analysis of a time series dataset select *Statistical methods* → *Seasonal Adjustment* → *Seasonal adjustment* → *Multiprocessing* → *New* and follow the ‘*JDemetra+ User Guide*’ (2016), case study 3.1.2.

Alternatively, the document can be created directly in the *Workspace* window. First, expand the *Seasonal adjustment* → *specifications*. Then, right click on the selected specification name under
tramoseats or x13 nodes then chose the option Create document. An empty document will be added to the relevant place in the Seasonal adjustment → documents section and the specification selected by the user will be assigned to it.

Figure 7.6: Opening a new document from the Seasonal adjustment → specifications level.

An empty document can also be created from the Seasonal adjustment → documents level. It can be done by selecting one of the items tramoseats, x13 or multi-documents and choosing an option New from the local menu (right click).

Figure 7.7: Creating a new document in the Documents section.

It is also possible to create a new document by choosing an option New from the local menu, which is available for the multi-documents item in the Workspace_#number tree.
All documents are added to the relevant part of the *Workspace* window. Once the workspace is saved (see 3.1) all the documents defined for this workspace are saved as well. The user may then investigate the saved results of the seasonal adjustment and update them following the *JDemetra+ User Guide* (2016), case study 3.1.1 (for single time series adjustment) or 3.1.2 (for seasonal adjustment of a dataset).

To use the documents created directly in the *Workspace* window, double click on its name to display it. To perform an analysis, drag and drop the series from the *Workspace* window to the document window (*SAProcessing-*\#number*, *TramoSeatsDoc-*\#number* or *X13Doc-*\#number*). For more details see items 3.1.1 and 3.1.2.
For the TramoSeatsDoc-#number and X13Doc-#number documents the seasonal adjustment processes is launched automatically and the results are displayed, once the user drags and drops series into the relevant place of the document.

In the case of the SAProcessing-#number document the user is expected to drag and drop the series into it and launch the process by clicking on the Start button (the green arrow marked in the picture below).

When the user clicks on an individual time series in the SAProcessing-#number window, the detailed results are displayed in the panel below the list of the series.
For each seasonal adjustment method the structure of the output is the same. The results of the seasonal adjustment process are divided into six parts: Input, Main results, Pre-processing, Decomposition, Benchmarking and Diagnostics. They are organised in a tree structure, which can be expanded. To investigate a specific node, click on it in the panel on the left to display its content in the panel on the right. Study the content of the section using the vertical scrollbar.

The structure and the content of the output presented in the Input and Pre-processing nodes is the same for both seasonal adjustment methods. It can be examined in 4.2. For the other nodes the content is specific to the decomposition step of the seasonal adjustment procedure. As both methods vary substantially in their approach to the decomposition, the description of the output is presented separately for both methods.
7.2.1. TramoSeats

The TRAMO/SEATS method consists of two linked programs: TRAMO and SEATS. The results from TRAMO, which are displayed under the Pre-processing node, are explained in 4.2. This section focuses on the nodes that are not handled there, in particular on the output produced by SEATS. Therefore, the sections that are explained here are: Main results, Decomposition, Benchmarking and Diagnostic. Their content is accessible once the user selects the appropriate node from the left-hand side seasonal adjustment results panel.

Figure 7.14: The structure of the results for TRAMO/SEATS.

7.2.1.1. Main results

On the very top of the Main results window, vital messages that concern the course of seasonal adjustment process are displayed (if any). They should be examined by the user as they indicate problems with the seasonal adjustment process that hinder the decomposition or even make it impossible. Some illustrative examples are given below.

When the time series is shorter than 3 years JDemetra+ is not able to perform the decomposition because of a lack of sufficient number of observations.

Errors

input: Not enough data

Figure 7.15: Error message displayed because of not enough available observations.

Although the procedure can be executed in the presence of outliers and missing values, their particular composition can stop the decomposition procedure. This is the case when missing values
or outliers cumulate in a specific period (e.g. the data gap is observed for every March) and therefore the results from the TRAMO model are not reliable enough to perform the decomposition.

**Errors**

*preprocessing: failed*

Figure 7.16: Error message due to concentration of missing values in a particular period.

In both cases discussed above the time series cannot be estimated by the TRAMO or RegARIMA models.

When the model chosen by TRAMO is further changed by SEATS, JDemetra+ displays the following warning.

**Warnings**

*decomposition.Model decomposition: Parameters cut off*

Figure 7.17: Information about the change of the ARIMA model made by SEATS.

Further sections of the *Main results* node includes summary information from TRAMO and SEATS and presents the main statistics that assess the quality of the outcomes.

Information about the TRAMO part of the process includes the estimation span, number of observations, transformation (if any) and deterministic effects. It is discussed in detail in 4.2.2.

**Pre-processing (Tramo)**

**Summary**

*Estimation span: [I-1995 - IV-2011]*
*68 observations*
*Series has been log-transformed*
*No trading days effects*
*No easter effect*
*1 detected outlier*

Figure 7.18: The summary of the results from TRAMO.

In the case of the pre-defined specifications RSA0, RSA1 and RSA3 no trading day effect is estimated. For the RSA2 and RSA4 pre-defined specifications, working day effects and the leap year effect are pre-tested and estimated if present. If the working day effect is significant, the pre-processing part includes the message *Working days effect (1 regressor)*. The message *Working days effect (2 regressors)* means that apart from the working day effect also the leap year effect has been estimated. For RSA5 and RSAfull the trading day effect and the leap year effect are pre-tested. If the trading day effect has been detected, either of the messages *Trading days effect (6 regressors)* or *Trading days effect (7 regressors)* are displayed, depending whether the leap year effect has been
detected or not. If the Easter effect is statistically significant, the message *Easter effect detected* is displayed.

The second part of the *Main results* panel presents the innovation variances of the orthogonal components estimated by SEATS which result from the decomposition of a stochastic time series (i.e. original series corrected for the deterministic effects and missing observations). These components are trend, seasonal, transitory and irregular and derive from the decomposition of the stochastic (linearized) time series $x_t$. In the case of additive decomposition $x_t$, can be presented as a sum of the components:

$$x_t = \sum_{i=1}^{k} x_{it}$$  \[5.1\]

where:

- $i$ - trend, seasonal, transitory or irregular components, respectively.

The linearised series $x_t$ follows an ARIMA model of the type:

$$\phi(B)\delta(B)x_t = \theta(B)a_t$$  \[5.2\]

where:

- $\delta(B)$ – a non-stationary autoregressive (AR) polynomial in $B$ (unit roots);
- $\theta(B)$ – an invertible moving average (MA) polynomial in $B$ which is a product of the invertible regular MA polynomial in $B$ and the invertible seasonal MA polynomial in $B^S$;
- $\phi(B)$ – a stationary autoregressive (AR) polynomial in $B$ and stationary seasonal polynomial in $B^S$;
- $a_t$ – a white-noise variable with the variance $V(a)$, also referred as *innovation*. Each component follows the general ARIMA model:

$$\phi_i(B)\delta_i(B)x_{it} = \theta_i(B)a_{it}$$  \[5.2\]

where:

- $a_{it} \sim WN(0,V(a))$ is an i.i.d. white-noise innovation of the $i^{th}$ component; it is also an estimator of the 1-period-ahead forecast error of the $i^{th}$ component.

---

68 Orthogonality means that behaviour of each component is uncorrelated with other components. In particular, causes of seasonal fluctuations are uncorrelated with causes of long term evolution of the series.

69 The MA process is invertible if can be rewritten as a linear combination of its past values plus the contemporaneous error term.

70 See 7.1.

71 It is assumed that irregular component is a white noise variable, which means that it follows ARIMA (0,0,0)(0,0,0) model.
The polynomials \( \theta_i(B) \), \( \phi_i(B) \) and \( \delta_i(B) \) are of finite order. A white-noise variable is normally, identically and independently distributed, with a zero-mean and variance of the component innovation (the variance of the 1-period ahead forecast error of the component) \( V(a) \). Two different components do not share the same unit autoregressive roots.

The components can be also expressed in compact form:

\[
\varphi_i(B)x_{it} = \theta_i(B)a_{it}
\]  \[5.3\]

where \( \varphi_i(B) \) is a product of the stationary \( \delta_i(B) \) and the non-stationary \( \phi_i(B) \) autoregressive polynomials.

SEATS decomposition fulfils the canonical property, that is it maximizes the variance of the irregular component providing trend, seasonal and transitory components as stable as possible (in accordance with the models)\(^{72}\). For each component the value of the innovation variance \( k_i \) is represented by the ratio of the component innovation variance \( V(a) \) to the component-ARIMA-model to variance of the series innovation \( V(a) \):\(^{73}\)

\[
k_i \frac{V(a)}{V(a)} = V(a_i)
\]  \[5.3\]

The innovations in the components are the cause of their stochastic behaviour (i.e., their moving features). The larger the variance, the more volatile the component will be\(^{74}\). In general, interest centres on more stable seasonal signals and hence when competing models are compared, the preferred one is the model that minimizes the innovation variance of the seasonal and trend components\(^{75}\). Such solution results in the trend and seasonal component that are as smooth and stable as possible, and the irregular component that absorb as much noise as possible.

---

\(^{72}\) In order to identify the components SEATS assumes that components are orthogonal to each other and each component except for the irregular one is clean of noise. This is called the canonical property, and implies that no additive white noise can be extracted from a component that is not the irregular one.

\(^{73}\) MARAVALL, A. (2009).


Figure 7.19: The summary of the results from SEATS.

The *Diagnostics* section is displayed below the summary of decomposition results produced by SETAS. It includes the outcome of the most important quality indicators. They are discussed in 5.2.1.4. The results are accompanied by two figures. The one on the left shows the original series, the final seasonally adjusted series, and the final trend.

![Graph showing original series, seasonally adjusted series, and trend](image)

**Figure 7.20: The original series and the results from decomposition: the seasonally adjusted series and the trend.**

The local menu, which can be activated by right clicking on the graph, contains the following functions:

- **Select all** – selects all time series presented in the graph;
- **Show title** – option is not currently available;
- **Show legend** – displays titles of all time series presented on the graph;
- **Edit format** – enables the user to change data format;
- **Color scheme** – allows the user to change the colour in the graph;
- **Lines thickness** – allows the user to choose between thin and thick lines to be used for a graph;
- **Show all** – this option is not currently available;
- **Export image to** – allows the graph to be sent to the printer and saved to the clipboard or as a file in PNG format;
- **Configure** – enables the user to customize chart and series display.

The panel on the right shows the SI ratios, which are discussed in detail in *JDemetra+ User Guide* (2016).

---

76 To be available in the *JDemetra+ User Guide* (2016).
The Main results three subsections: Charts, Table and S-I ratio node provide visual presentation of the decomposition results. The Charts section includes a chart of the original series, trend and seasonally adjusted series together with the forecasts for the next year for each of these series.

The second graph from the Charts section presents the calendar effects, the seasonal component and the irregular component. As a rule, the calendar component is expected to be weaker than the seasonal component; however it is not the case for non-seasonal series with calendar effects present. The irregular component is assumed to be random and unpredictable; therefore its forecasts are, in general, zero (for the additive model) or one (for the multiplicative model). The lack of certain movements (seasonal and/or irregular) is manifested by a horizontal line with values equal to zero (for the additive model) or one (for the multiplicative model).
For both charts the local menu, activated by right click on the graph, offers the same set of options as the ones available for the chart presented in Figure 7.20. The Table section presents the data for the original series with forecasts, the final seasonally adjusted series, the trend with forecasts, the seasonal component with forecasts and the irregular component. In general, the irregular component by definition cannot be forecast, therefore for TRAMO/SEATS its forecasts are set to 0 (additive model) or 1 (multiplicative model). However, there are some exceptions from this rule. Firstly, when the transitory component was extracted, it is not displayed in a separate table, but together with the irregular component. In such case the forecasts displayed in the Irregular column result from forecasting the transitory component. Secondly, non-zero irregular forecasts can also be explaining by TC outliers (especially when they were identified at the end of the series), as this outlier generates the effect that decays in time\textsuperscript{77}. Finally, in the case of log models, SEATS performs a bias correction to ensure that the irregular and seasonal components are, on average, equal to 1. The correction factor is also applied to the forecasts, which are then not exactly 1 (but they are constant). When the X-13ARIMA-SEATS method is used no values for the forecasting period of the irregular component are computed, following the approach implemented in the original software.

The local menu, activated by right clicking on the table, contains the same functions that are available for Grid (see description below Figure 3.22). All series are extended with one year of forecasts. These forecasts are presented in the bottom part of the table.

\textsuperscript{77} See 7.1.1.
The S-I ratio chart presents the seasonal-irregular (S-I) component and the seasonal factors for each of the periods in the time series (months or quarters). The seasonal-irregular component is calculated as the ratio of the original series to the estimated trend, thus it presents an estimate of the detrended series. Blue curves represent the final seasonal factors and the red straight lines represent the mean seasonal factor for each period. For the investigation of the S-I ratio tool and options available for this chart refer to the ‘JDemetra+ User Guide’ (2016).

7.2.1.2. Decomposition

SEATS receives the linearised series from TRAMO. The decomposition made by SEATS assumes that all components in time series - trend, seasonal, transitory (if present) and irregular - are orthogonal and can be expressed by the ARIMA model. SEATS performs the canonical decomposition of the components which assumes that only irregular components include noise.
Each model is presented in closed form (i.e. using the backshift operator $B$). In the main page of the Decomposition node the following items are presented (if estimated):

- **Model** – the ARIMA model for the linearised series and the innovation variance of the model for the linearised series;
- **$Sa$** – the ARIMA model for the seasonally adjusted series and the innovation variance of the model for the seasonally adjusted series;
- **Trend** – the ARIMA model for the trend component of the series and the innovation variance of the model for the trend;
- **Seasonal** – the ARIMA model for the seasonal component of the series and the innovation variance of the model for the seasonal component;
- **Transitory** – the ARIMA model for the transitory component of the series and the innovation variance of the model for the transitory component;
- **Irregular** – The innovation variance of the irregular component.

The trend component captures the low-frequency variation of the series and displays a spectral peak at frequency 0. By contrast, the seasonal component picks up the spectral peaks at seasonal frequencies and the irregular component captures a white noise process. The transitory component, which is the additional component estimated by SEATS for some time series, can be seen as a noise-free, detrended and seasonally adjusted time series. This component captures highly transitory variation that is not white noise and should not be assigned to the seasonal component or the trend component. It will capture spectral peaks at frequencies that are neither zero nor seasonal\(^{78}\). The model for the transitory component is a stationary ARMA model, with low-order MA components (order $Q - P$, when $Q > P$)\(^{79}\) and AR roots with small moduli that should not be included in the trend component or the seasonal component.

The example of the time series decomposition calculated by SEATS is presented below. It can be seen that the overall autoregressive polynomial has been factorized into the polynomials assigned to the components according to the frequencies of the roots. As an example, the model for the trend is: $\text{ARIMA}(1,2,3)(0,0,0)$ with innovation variance 0.22517.

---


\(^{79}\) Q is an order of the MA process, while P is an order of the AR process.
The decomposition above is performed for the ARIMA model identified by TRAMO. However, in some cases, the ARIMA model chosen by TRAMO is changed by SEATS\(^8\). It is done, for example, when the ARIMA model selected by TRAMO leads to a non-admissible decomposition. In this case, the ARIMA model chosen by SEATS is displayed in the Pre-processing → Arima section of the result panel and this model is used to decompose the series.

In the Pre-processing → Arima section two spectra are presented: one from the TRAMO model and one from the SEATS model.

\(^8\) See 7.1.2.1 for further information.
Under the section where the polynomials and the regular AR roots from the TRAMO model are reported, the respective information is also displayed for the model identified by SEATS.

Tramo model $(1,1,0)(1,0,0)$

**Polynomials**
- regular AR: $1.0000 - 0.576367 B$
- seasonal AR: $1.0000 - 0.373660 S$

**Regular AR inverse roots**
- argument=0.0000, modulus=0.5734

Seats model $(1,1,0)(0,1,1)$

**Polynomials**
- regular AR: $1.0000 - 0.569729 B$
- seasonal MA: $1.0000 - 0.601211 S$

**Regular AR inverse roots**
- argument=0.0000, modulus=0.5807

7.2.1.2.1. Stochastic series

A discrete-time stochastic process is a collection of random variables $\{X_t(\omega)\}$, where $t$ denotes time and $\omega$ denotes an elementary event. A time series associated with these random variables is called a stochastic time series. In general, a stochastic time series is made of two components, one which is predictable once the history of the process $X_{t-1}$ is known, and one that is not.

This part of the output presents the results of the decomposition of the stochastic series resulting from linearization procedure performed by TRAMO. The main panel incorporates the table containing the following series produced by SEATS:

- Series (lin);
- Seasonally adjusted (lin);
- Trend (lin);
- Seasonal (lin);
- Irregular (lin);
- Seasonally adjusted (stde lin);
- Trend (stde lin); Seasonal (stde lin);
Irregular (stde lin) (contains transitory component, if any).

Lin is an abbreviation from linearised series (including logarithmic transformation of the data if the multiplicative decomposition is used) and stde denotes standard deviation. All series are extended with one year of forecasts. These forecasts are presented in the bottom part of the table.

<table>
<thead>
<tr>
<th></th>
<th>Series (ln)</th>
<th>Seasonally adjusted (ln)</th>
<th>Trend (ln)</th>
<th>Seasonal (ln)</th>
<th>Irregular (ln)</th>
<th>Seasonally adjusted (stde ln)</th>
<th>Trend (stde lin)</th>
<th>Seasonal (stde ln)</th>
<th>Irregular (stde ln)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1995</td>
<td>72.776</td>
<td>72.777</td>
<td>72.77</td>
<td>0.005</td>
<td>0.061</td>
<td>0.355</td>
<td>0.359</td>
<td>0.355</td>
<td>0.690</td>
</tr>
<tr>
<td>1996</td>
<td>72.076</td>
<td>71.155</td>
<td>71.169</td>
<td>-0.075</td>
<td>-0.014</td>
<td>0.317</td>
<td>0.311</td>
<td>0.317</td>
<td>0.653</td>
</tr>
<tr>
<td>1997</td>
<td>72.476</td>
<td>73.613</td>
<td>73.193</td>
<td>-0.137</td>
<td>0.062</td>
<td>0.363</td>
<td>0.363</td>
<td>0.363</td>
<td>0.662</td>
</tr>
<tr>
<td>1998</td>
<td>72.476</td>
<td>73.613</td>
<td>73.193</td>
<td>-0.137</td>
<td>0.062</td>
<td>0.363</td>
<td>0.363</td>
<td>0.363</td>
<td>0.662</td>
</tr>
<tr>
<td>1999</td>
<td>74.476</td>
<td>74.591</td>
<td>74.481</td>
<td>-0.025</td>
<td>0.019</td>
<td>0.253</td>
<td>0.249</td>
<td>0.253</td>
<td>0.692</td>
</tr>
<tr>
<td>2000</td>
<td>75.076</td>
<td>74.876</td>
<td>74.830</td>
<td>0.021</td>
<td>0.036</td>
<td>0.248</td>
<td>0.248</td>
<td>0.248</td>
<td>0.692</td>
</tr>
<tr>
<td>2001</td>
<td>75.776</td>
<td>75.049</td>
<td>75.082</td>
<td>0.727</td>
<td>-0.023</td>
<td>0.248</td>
<td>0.244</td>
<td>0.248</td>
<td>0.692</td>
</tr>
<tr>
<td>2002</td>
<td>75.776</td>
<td>75.397</td>
<td>75.398</td>
<td>0.378</td>
<td>-0.061</td>
<td>0.248</td>
<td>0.244</td>
<td>0.248</td>
<td>0.692</td>
</tr>
<tr>
<td>2003</td>
<td>75.776</td>
<td>75.397</td>
<td>75.398</td>
<td>0.378</td>
<td>-0.061</td>
<td>0.248</td>
<td>0.244</td>
<td>0.248</td>
<td>0.692</td>
</tr>
<tr>
<td>2004</td>
<td>75.776</td>
<td>75.397</td>
<td>75.398</td>
<td>0.378</td>
<td>-0.061</td>
<td>0.248</td>
<td>0.244</td>
<td>0.248</td>
<td>0.692</td>
</tr>
<tr>
<td>2005</td>
<td>75.776</td>
<td>75.397</td>
<td>75.398</td>
<td>0.378</td>
<td>-0.061</td>
<td>0.248</td>
<td>0.244</td>
<td>0.248</td>
<td>0.692</td>
</tr>
<tr>
<td>2006</td>
<td>75.776</td>
<td>75.397</td>
<td>75.398</td>
<td>0.378</td>
<td>-0.061</td>
<td>0.248</td>
<td>0.244</td>
<td>0.248</td>
<td>0.692</td>
</tr>
<tr>
<td>2007</td>
<td>75.776</td>
<td>75.397</td>
<td>75.398</td>
<td>0.378</td>
<td>-0.061</td>
<td>0.248</td>
<td>0.244</td>
<td>0.248</td>
<td>0.692</td>
</tr>
<tr>
<td>2008</td>
<td>75.776</td>
<td>75.397</td>
<td>75.398</td>
<td>0.378</td>
<td>-0.061</td>
<td>0.248</td>
<td>0.244</td>
<td>0.248</td>
<td>0.692</td>
</tr>
<tr>
<td>2009</td>
<td>75.776</td>
<td>75.397</td>
<td>75.398</td>
<td>0.378</td>
<td>-0.061</td>
<td>0.248</td>
<td>0.244</td>
<td>0.248</td>
<td>0.692</td>
</tr>
<tr>
<td>2010</td>
<td>75.776</td>
<td>75.397</td>
<td>75.398</td>
<td>0.378</td>
<td>-0.061</td>
<td>0.248</td>
<td>0.244</td>
<td>0.248</td>
<td>0.692</td>
</tr>
<tr>
<td>2011</td>
<td>75.776</td>
<td>75.397</td>
<td>75.398</td>
<td>0.378</td>
<td>-0.061</td>
<td>0.248</td>
<td>0.244</td>
<td>0.248</td>
<td>0.692</td>
</tr>
<tr>
<td>2012</td>
<td>75.776</td>
<td>75.397</td>
<td>75.398</td>
<td>0.378</td>
<td>-0.061</td>
<td>0.248</td>
<td>0.244</td>
<td>0.248</td>
<td>0.692</td>
</tr>
<tr>
<td>2013</td>
<td>75.776</td>
<td>75.397</td>
<td>75.398</td>
<td>0.378</td>
<td>-0.061</td>
<td>0.248</td>
<td>0.244</td>
<td>0.248</td>
<td>0.692</td>
</tr>
<tr>
<td>2014</td>
<td>75.776</td>
<td>75.397</td>
<td>75.398</td>
<td>0.378</td>
<td>-0.061</td>
<td>0.248</td>
<td>0.244</td>
<td>0.248</td>
<td>0.692</td>
</tr>
</tbody>
</table>

Figure 7.29: Stochastic series extended with forecasts.

The two subsections allow for some insights into the development of the two components (trend and seasonal) in the last 7 years and a forecast for the next year. For each component its values are displayed with the associated 95% confidence intervals, highlighting the fact that these values result from the estimation procedure. The width of the confidence intervals shows the size of uncertainty of the estimation results, which in general is greater at the end of the time series. The prediction intervals, which are available for the forecasts, are even wider than the confidence intervals.

The graph is available for the trend (figure 5.30) and for the seasonal component (figure 5.31).

Figure 7.30: The trend estimate with the confidence interval and the prediction interval.
7.2.1.2.2. Components

The Components section presents the values of the components calculated from the ARIMA model applied to the linearised series. Cmp is an abbreviation from component.

<table>
<thead>
<tr>
<th>Year</th>
<th>Series (cmp)</th>
<th>Seasonally adjusted (cmp)</th>
<th>Trend (cmp)</th>
<th>Seasonal (cmp)</th>
<th>Irregular (cmp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-2000</td>
<td>32,531</td>
<td>36,812</td>
<td>36,009</td>
<td>0.281</td>
<td>1</td>
</tr>
<tr>
<td>2-2000</td>
<td>25,423</td>
<td>36,549</td>
<td>36,9</td>
<td>0.798</td>
<td>0.999</td>
</tr>
<tr>
<td>3-2000</td>
<td>37,233</td>
<td>37,069</td>
<td>36,998</td>
<td>1.024</td>
<td>1.002</td>
</tr>
<tr>
<td>4-2000</td>
<td>38,149</td>
<td>37,065</td>
<td>37,099</td>
<td>1.029</td>
<td>0.999</td>
</tr>
<tr>
<td>5-2000</td>
<td>34,887</td>
<td>37,211</td>
<td>37,205</td>
<td>0.938</td>
<td>1</td>
</tr>
<tr>
<td>6-2000</td>
<td>35,048</td>
<td>37,283</td>
<td>37,32</td>
<td>1.047</td>
<td>0.999</td>
</tr>
<tr>
<td>7-2000</td>
<td>38,331</td>
<td>37,444</td>
<td>37,443</td>
<td>1.024</td>
<td>1</td>
</tr>
<tr>
<td>8-2000</td>
<td>40,487</td>
<td>37,603</td>
<td>37,573</td>
<td>1.077</td>
<td>1.001</td>
</tr>
<tr>
<td>9-2000</td>
<td>37,448</td>
<td>37,773</td>
<td>37,706</td>
<td>0.994</td>
<td>1.002</td>
</tr>
<tr>
<td>10-2002</td>
<td>41,869</td>
<td>37,851</td>
<td>37,84</td>
<td>1.098</td>
<td>1</td>
</tr>
<tr>
<td>11-2000</td>
<td>42,486</td>
<td>38,017</td>
<td>37,979</td>
<td>1.118</td>
<td>1.001</td>
</tr>
<tr>
<td>12-2000</td>
<td>35,706</td>
<td>38,083</td>
<td>38,127</td>
<td>1.016</td>
<td>0.999</td>
</tr>
<tr>
<td>1-2001</td>
<td>34,504</td>
<td>38,248</td>
<td>38,288</td>
<td>0.902</td>
<td>0.999</td>
</tr>
</tbody>
</table>

Figure 7.32: The theoretical components calculated from the ARIMA model applied to the linearized series.

7.2.1.2.3. WK analysis – Components

This section presents the (pseudo) spectra of the components and seasonally adjusted series calculated from the ARIMA models presented in the main panel of the Decomposition section. The sum of the spectra of the components should be equal to the spectrum of the linearised time series, which is presented in the Pre-processing → Arima node. When the TRAMO model has not been accepted by SEATS, the spectra of the components derive from the ARIMA model changed by SEATS. The spectrum of the seasonally adjusted series (yellow) is the sum of the spectra of the trend component (green), the transitory component, if present (black), and the irregular component (orange).
The stochastic variability in the $i^{th}$ component is generated by the innovations $a_t$. Therefore, small values of variance innovations $V(a_i)$ are associated with a stable component, and large values of $V(a_i)$ with an unstable component. The spectrum of the $i^{th}$ component is proportional to $V(a_i)$. Further, stable trend and seasonal components are those with thin spectral peaks, while unstable ones are characterised by wide spectral peaks.

If a series contains an important component for a certain frequency, its spectrum should reveal a peak around that frequency. As a trend can be thought as a cyclical component with an infinite length of the movement, the spectral peak should occur at the frequency $\omega = 0$. For a monthly time series there are six seasonal frequencies: $\frac{\pi}{6}, \frac{2\pi}{6}, \frac{3\pi}{6}, \frac{4\pi}{6}, \frac{5\pi}{6}, \pi$, while for quarterly data there are two seasonal frequencies: $\frac{\pi}{2}, \pi$. The spectrum for a seasonal component has peaks around these frequencies.

In Figure 7.33 standard spectra for trend (green), seasonal (blue), transitory (black) and irregular (orange) are displayed. As it was already explained, the frequency $\omega = 0$ is associated with a trend. For the frequencies in the range $[0 + \epsilon_1, \pi - \epsilon_2]$ with $\epsilon_1, \epsilon_2 > 0$ and $\epsilon_1 < \pi - \epsilon_2$ the associated period $\frac{\pi}{2}$ will be longer than a year and bounded. The frequencies in the range $[0, \frac{\pi}{2}]$ are associated with short term movements, with a cycle completed in less than a year. Finally, the peak in the transitory component is evidence of a trading day effect.

Figure 7.33: The theoretical spectra of the components and the seasonally adjusted series.

---

82 Ibid.
The ACGF (stationary) window displays the pseudo-autocovariance generating functions\textsuperscript{83} \textsuperscript{84} of the stationary components. They are theoretical values (i.e. they are not computed on the linearised series, but on the ARIMA model).

![Figure 7.34: ACGF function of the stationary components.](image)

\textbf{7.2.1.2.4. WK analysis – Final estimators}

The various graphs that can be found in this section show the results of the estimation of the components performed with Wiener-Kolmogorov (WK) filters\textsuperscript{91}. The Wiener-Kolmogorov filters are symmetric and bi-infinite, which requires that for each observation the past and future observations exist. Obviously, they are not available at the beginning and end of the time series. In order to apply a filter to all observations from $x_t$, the original, linearised time series is extended with forecasts and backcasts using an ARIMA model, which has been chosen in the TRAMO phase of seasonal adjustment. Then, SEATS applies the filter to extended series\textsuperscript{85}. As a new observation (i.e. observation for period $t + 1$) is available, the forecast for period $t + 1$ is replaced by this new observation and all forecasts for periods $g > t + 1$ are updated. It means that at the end of the time series the estimator of the component is preliminary and is subject to revisions, while in the central periods the estimator can be treated as final (also called a historical estimator)\textsuperscript{86}.

Regarding the importance of final (historical) estimators derived applying the WK filters, JDemetra+ presents several graphs showing their properties (see 7.1.2.2). The corresponding graphs for the components and for the final estimators of the components vary, as components and final estimators follow different models. For example, the seasonal component follows the model:

\textsuperscript{83} See 7.3.
\textsuperscript{84} See 7.1.2.2.\textsuperscript{91}
\textsuperscript{85} See 7.1.2.2.
\textsuperscript{86} CLEVELAND, W.P., and TIAO, C.G. (1976).
\textsuperscript{86} MARAVALI, A. (2011).
\[ \phi_s(B)s_t = \theta_s(B)a_t, \] 
while the corresponding MMSE estimator of the seasonal component follows the model: 
\[ \phi_s(B)\hat{s}_t = \theta_s(B)\alpha_s(F)a_t, \] 
where \( \alpha_s(F) \) is a difference between the theoretical component and the estimator.\(^{87}\)

These graphs are listed below. For each graph all series from the graph can be copied (the Copy all visible). The graph can be saved and/or printed. These options are available from the local menu, which is available by right clicking on the graph.

**Spectrum**

The spectra\(^{88}\) of the final estimators are shown in the first graph. The spectrum of the estimator of the seasonal component is obtained by multiplying the squared gain\(^{89}\) of the filter by the spectrum of the linearised series.

From the example below it is clear that these spectra are similar to those of the components, although estimator spectra show spectral zeros at the frequencies where the component spectra are close but not exactly zero. The estimator adapts to the structure of the analysed series, i.e. the width of the spectral holes in the seasonally adjusted series (yellow line) depends on the width of the seasonal peaks in the seasonal component estimator spectrum (blue lines).\(^{90}\)

![Figure 7.35: Spectra of the final estimators.](image)

**Square gain of components filter**
The squared gain controls the extent to which a movement of particular amplitude at a frequency \( \omega \) is delivered to the output series. It determines how the variance of the series contributes to the variance of the component for the different frequencies. In other words, it specifies which

---

\(^{87}\) MARAVALL, A. (2008b).

\(^{88}\) See 7.3.

\(^{89}\) See 7.1.2.2.

\(^{90}\) MARAVALL, A. (2003).
frequencies will contribute to the signal (that is, it filters the spectrum of the series by frequencies)\textsuperscript{91}. When the squared gain is zero in band $[\omega_1, \omega_2]$ it means that the output series is free of movements in this range of frequencies\textsuperscript{92}. On the contrary, when for some $\omega$ the squared gain is 1, then all variation is passed on to the component estimator.

Figure 5.36 shows that the seasonal frequencies are assigned to the seasonal component while the seasonally adjusted series captures the non-seasonal frequencies. As a consequence, it is expected that the seasonal component estimator captures only the seasonal frequencies, so its peaks assume unitary values at seasonal frequencies. By contrast, the estimator of the non-seasonal part of the time series is expected to eliminate seasonal frequencies, leaving unmodified non-seasonal frequencies. Therefore, the squared gain of seasonally adjusted data should be nearly zero for seasonal frequencies.

The squared gain shape depends on the model used for the time series. Figures 5.36 and 5.37 show squared gains derived from two different models are represented. In figure 5.36, the squared gain of the seasonal adjustment filter shows relatively large troughs to suppress a highly stochastic seasonal component. In contrast, figure 5.37 presents the estimators for a seasonal component which is much more deterministic than the previous one.

---

\textsuperscript{91} Squared gain definition is in 7.1.2.2.

\textsuperscript{92} PLANAS, C. (1998).
WK filters The Wiener-Kolmogorov (WK) filter \( v_i(B, F) \) shows the weights that have been applied to the original series \( x_t \) to estimate the \( \hat{x}_{it} \) component in the following way (see 7.1.2.2 for a description of the WK filter):

\[
\hat{x}_{it} = v_i(B, F)x_{it}
\]  \[5.4\]

where:

\[
v_i(B, F) = v_0 + \sum_{j=0}^{\infty} (B^j + F^j) \]  \[5.5\] Since each WK filter is symmetric and centred, it is also convergent, which enables the user to approximate an infinite number of realizations \( x_t \) by a finite number of terms (from the graph below it could be noticed that \( j = 36 \), so the WK filter includes \( 36 + 1 + 36 = 73 \) terms).

Since WK filters are symmetric, centered and convergent, they are valid for computing the estimators in the central periods of the sample. The following graph demonstrates the weights that are applied to each observation for each component to calculate the estimate of each component.
Figure 7.38: WK filter weights for the seasonally adjusted series and for the components.

The highest weights are applied to the central observations and the weights decrease for observations further away in time. Therefore the estimated value of each component is highly influenced by the linearised series value. The weighting pattern depends on the component. For example, in the case of the seasonal component the greatest weights are applied to the current value and the past and future values from this same period. On the contrary, the estimate of the trend at a given point in time is mostly influenced by the current value and few preceding and following values of the linearised series.

PsiE-weights

PsiE-weights ($\psi$) are a different representation of the final estimator, i.e. this representation shows the estimator as a filter applied to the innovations $a_t$, rather than the series $x_t$. Figure 7.39 shows for each component how the contribution of the total innovation to the component estimator $\hat{x}_t$ varies in time (the size of this contribution is shown on the Y-axis). For non-negative values on the X-axis, PsiE-weights show the effect of starting conditions, present and past innovations in series, while for negative observations they present the effect of future innovations. It can be seen that they are non-convergent in the past (they are convergent when series $x_t$ is stationary). On the contrary, the effect of future innovations is a zero-mean and convergent process. PsiE-weights are important to analyse the convergence of estimators and revision errors.

See 7.1.2.3. For further details see MARAVALL, A. (2008b).
7.2.1.2.5. Error analysis

Being obtained by using forecasts, the component estimators at the end points of the series are preliminary and are subject to revisions as future data become available, until it can be assumed that the final (historical) estimator has been reached. This process typically last between 3 and 5 years\(^{94}\). The error analysis deals with the size of error variances and speed of their convergence to the final value.

From the model-based structure it is possible to determine the underlying models for the two types of errors: revision error and estimation error. Thus the respective variances, autocorrelations and spectra can be computed. The speed of convergence of the revision can also be assessed. For each \(i^{th}\) component, the total error in the preliminary estimator \(d_{it|t+k}\) is expressed as:

\[
d_{it|t+k} = x_{it} - \hat{x}_{it|t+k}
\]  \[5.6\]

where:

- \(x_{it}\) is the \(i^{th}\) component;
- \(\hat{x}_{i|t+k}\) is the estimator of \(x_{it}\) when the last observation is \(x_{t+k}\) (\(x_{it}\) is a time series).

The total error in the preliminary estimator can be also presented as a sum of the final estimation error \((e_{it})\) and the revision error \((r_{it|t+k})\), i.e.:

\[
d_{it|t+k} = x_{it} - \hat{x}_{it|t+k} = (x_{it} - \hat{x}_{it}) + (\hat{x}_{it} - \hat{x}_{it|t+k}) = e_{it} + r_{it|t+k} \]  \[5.7\]

The final estimation error \((e_{it})\) and the revision error \((r_{it|t+k})\) are orthogonal\(^{95}\).


\(^{95}\) MARAVALL, A. (2000).
The errors analysis (variance and ACF of the total error and its components) is available for the trend and for the seasonally adjusted series. The values are given in units of variance of the innovations in the linearized series. Therefore, from the examples presented below it can be noticed that the variance of the concurrent estimator of the seasonally adjusted series is roughly 17% of the variance of the innovations of the linearized series (Figure 7.40) and 27% for the trend (Figure 7.41).

<table>
<thead>
<tr>
<th>Lag</th>
<th>Final error</th>
<th>Revision error (concurrent estimator)</th>
<th>Total error (concurrent estimator)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.3740</td>
<td>0.0370</td>
<td>0.4042</td>
</tr>
<tr>
<td>2</td>
<td>0.1763</td>
<td>0.2247</td>
<td>0.2009</td>
</tr>
<tr>
<td>3</td>
<td>-0.6745</td>
<td>-0.0172</td>
<td>-0.0453</td>
</tr>
<tr>
<td>4</td>
<td>-0.2282</td>
<td>-0.1817</td>
<td>-0.2045</td>
</tr>
<tr>
<td>5</td>
<td>-0.5295</td>
<td>-0.2970</td>
<td>-0.3134</td>
</tr>
<tr>
<td>6</td>
<td>-0.3673</td>
<td>-0.3554</td>
<td>-0.3612</td>
</tr>
<tr>
<td>7</td>
<td>-0.3473</td>
<td>-0.3559</td>
<td>-0.3517</td>
</tr>
<tr>
<td>8</td>
<td>-0.2675</td>
<td>-0.2949</td>
<td>-0.2814</td>
</tr>
<tr>
<td>9</td>
<td>-0.1386</td>
<td>-0.1798</td>
<td>-0.1596</td>
</tr>
<tr>
<td>10</td>
<td>0.0693</td>
<td>0.0212</td>
<td>0.0448</td>
</tr>
<tr>
<td>11</td>
<td>0.2436</td>
<td>0.2026</td>
<td>0.2228</td>
</tr>
<tr>
<td>12</td>
<td>0.7891</td>
<td>0.7405</td>
<td>0.7644</td>
</tr>
</tbody>
</table>

Figure 7.40: Autocorrelation of the errors for the seasonally adjusted series – the case of deterministic seasonal pattern.

<table>
<thead>
<tr>
<th>Lag</th>
<th>Final error</th>
<th>Revision error (concurrent estimator)</th>
<th>Total error (concurrent estimator)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.6436</td>
<td>0.1490</td>
<td>0.2093</td>
</tr>
<tr>
<td>2</td>
<td>0.1708</td>
<td>0.2535</td>
<td>0.2140</td>
</tr>
<tr>
<td>3</td>
<td>-0.0089</td>
<td>0.0813</td>
<td>0.0434</td>
</tr>
<tr>
<td>4</td>
<td>-0.1150</td>
<td>-0.0499</td>
<td>-0.0763</td>
</tr>
<tr>
<td>5</td>
<td>-0.1806</td>
<td>-0.1384</td>
<td>-0.1596</td>
</tr>
<tr>
<td>6</td>
<td>-0.2066</td>
<td>-0.2033</td>
<td>-0.2033</td>
</tr>
<tr>
<td>7</td>
<td>-0.1940</td>
<td>-0.2255</td>
<td>-0.2164</td>
</tr>
<tr>
<td>8</td>
<td>-0.1438</td>
<td>-0.2133</td>
<td>-0.1501</td>
</tr>
<tr>
<td>9</td>
<td>-0.0589</td>
<td>-0.1627</td>
<td>-0.1122</td>
</tr>
<tr>
<td>10</td>
<td>0.0613</td>
<td>-0.0692</td>
<td>-0.0068</td>
</tr>
<tr>
<td>11</td>
<td>0.1809</td>
<td>0.0947</td>
<td>0.1359</td>
</tr>
<tr>
<td>12</td>
<td>0.2396</td>
<td>0.2640</td>
<td>0.2523</td>
</tr>
</tbody>
</table>

Figure 7.41: Autocorrelation of the errors for the trend – the case of a non-seasonal series.

As stressed by MARAVALL, A. (1995), large revisions are associated with highly stochastic components and converge quickly, while smaller revisions are implied by very stable components and converge slowly. In general, the slow convergence of the SA series estimator to the final estimator suggests that very little would be gained from moving from a current adjustment to a
concurrent one. When several models for a given series are compared, the preferred is the one for which the historical estimation error is minimal and convergence of revision errors is relatively fast, even if the resulting ARIMA model is optimal from the TRAMO perspective.

The last table available in the Errors analysis panel presents the convergence of a concurrent estimator measured by the revision error, which is the difference between preliminary and final estimator. For each component the table shows the percentage reduction in the standard error of the revision after an additional year (up to 5 years). Comparisons are made with the concurrent estimators. This table informs about the time needed by the concurrent estimators to converge to the final ones and therefore how many periods it takes for a new observation to no longer significantly affect the estimate.

For the series used to produce Figure 7.40, Figure 7.41 and Figure 7.42 the estimated innovation variance of the seasonal component is 0.02135 which implies that this component is fairly stable, while the trend component, for which the innovation variance is 0.11526, is more stochastic. The convergence of the SA series is slow (Figure 7.42), so in this case the current adjustment strategy would imply a little loss in precision of the SA series. It is because the stable component is of little importance in explaining the series variability; therefore the removal of this component implies the smaller revision error, which tends to converge slowly.

Revised errors

| Percentage reduction in the standard error of the revision after additional years (comparison with concurrent estimators) |
|--------------------------------------------------|---------------|---------------|---------------|---------------|---------------|
| After... | 1 year | 2 years | 3 years | 4 years | 5 years |
| sa       | 25.94% | 45.52% | 59.92% | 70.52% | 76.31% |
| trend    | 59.64% | 70.31% | 78.16% | 83.93% | 88.18% |

Figure 7.42: Revision errors analysis up to five years of observations.

When a series is non-seasonal and a non-seasonal ARIMA model is used, then the autocorrelation errors table for seasonally adjusted data is not displayed. In this case the final error, the revision error and the total error, displayed for the trend component, converge to zero swiftly.

---

96 GÓMEZ, V., and MARAVALL, A. (2001). The current adjustment means that model, filters, outliers and regression parameters are re-identified and the respective parameters and factors re-estimated at appropriately set review periods (usually once a year). The seasonal and calendar factors to be used to adjust for seasonal and calendar effects of new unadjusted data in-between the review periods are those estimated in the previous review period and forecasted up to the next review period. In the concurrent adjustment strategy the model, filters, outliers, regression parameters are re-identified and the respective parameters and factors re-estimated every time new or revised data become available. Source: 'ESS Guidelines on Seasonal Adjustment' (2015).

7.2.1.2.6. Growth rates

The tables in this section show, starting with concurrent estimation, the convergence of the trend and seasonally adjusted series to their final estimators as new observations become available. The calculation is performed for the growth rates of the time series $z_t$ over the period $(t - m, t)$. The error variances displayed in this section are based on the estimation error of the stochastic trend and the seasonally adjusted series. The errors in the parameter estimates are not considered.

When the multiplicative model is applied, the growth rate over the $m$ periods is defined as

$$\frac{Z_t}{Z_{t-t+m}} - 1 \times 100.$$

All standard errors reported for the growth rates in the following tables are computed using a linear approximation to the rates. When period-to-period changes are large, these standard errors should be interpreted as broad approximations, that will tend to underestimate the true values. In the case of the additive model the growth rate is given by the difference between $Z_{t-t+m}$ and $Z_t$.

The total estimation error is the largest for the first period (concurrent estimator) and it decreases (for preliminary estimators) until it reaches a constant value. This constant value is the standard deviation of the historic estimator. In Figure 7.44 the total estimation error for historic estimator is around 0.38. From the observation of convergence, it can be judged that after 6 periods a new observation no longer significantly affects the estimate (the standard error is 0.38% from this point onwards).
Figure 7.44: Analysis of growth rates for seasonally adjusted series.

For the series presented in Figure 7.44 and Figure 7.45 the trend estimator converges faster than that of the seasonally adjusted series because the trend component is stochastic while the seasonal component is rather stable. After 2 years (8 additional observations) all estimators have practically converged (revision error is close to zero).

Figure 7.45: Analysis of growth rates for the trend.

7.2.1.2.7. Model based tests

The Model-based tests section concentrates on the distribution of components, theoretical estimators and empirical estimates (stationary transformation). This node is divided into three sections.

Variance

In this panel the variances of the component innovations are displayed (variance of the component innovation, (Component), theoretical variances of the stationary transformation of the estimated

---


99 The variance of the component innovation is the variance resulting from the ARIMA model for this component.
components and empirical variances of the stationary transformation of the estimated components \((\text{Estimate})\) are displayed\(^{100}\) (see also section 5.2.2.1).

SEATS identifies the components assuming that, except for the irregular component, they are clean of noise. It implies that the variance of the irregular is maximized in contrast to the trend and seasonal components which are as stable as possible. In this section JDemetra+ presents a table, which compares the variance of the stationary transformation of the innovations of the components \((\text{Component})\) with the variance of their theoretical estimators \((\text{Estimator})\) and the variance of their empirical (actually obtained) estimates \((\text{Estimate})\).

<table>
<thead>
<tr>
<th>Variance</th>
<th>Component</th>
<th>Estimator</th>
<th>Estimate</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trend</td>
<td>0.7020</td>
<td>0.5877</td>
<td>0.5233</td>
<td>0.7288</td>
</tr>
<tr>
<td>Seasonally adjusted</td>
<td>0.6785</td>
<td>0.7309</td>
<td>0.6795</td>
<td>0.6176</td>
</tr>
<tr>
<td>Seasonal</td>
<td>0.1097</td>
<td>0.0063</td>
<td>0.0034</td>
<td>0.2788</td>
</tr>
<tr>
<td>Irregular</td>
<td>0.0889</td>
<td>0.0266</td>
<td>0.0232</td>
<td>0.6354</td>
</tr>
</tbody>
</table>

Figure 7.46: Variance of the components, their theoretical estimators and their empirical estimates (stationary transformation).

It follows from the properties of the MMSE (Minimum Mean Square Error) estimator that it will always underestimate the component (estimators always have a smaller variance than the components). The size of underestimation depends on the particular model. The underestimation is relatively large when the variance of the component is relatively small. It means that, for example, the trend estimator always has a smaller variance than the trend component and the ratio of the two variances increases as the trend becomes more stable. Therefore, the more stochastic the trend is, the closer the estimator variance is to the component variance. On the other hand, the variance of the estimator of a very stable trend will be considerably lower than the variance of the component\(^{101}\). It means that the trend estimator provides a more stable trend than the one implied by the theoretical model\(^{102}\).

For all components it is expected that\(^{110}\) the variance of the component innovation is greater than the variance of its theoretical estimators\(^{103}\). The later should be relatively close to the variance of the empirical estimate\(^{104}\). As a rule, if the correct WK filter was applied then the empirical variance

\(^{100}\) MARAVALL, A. (1995).


\(^{103}\) From the TRAMO/SEATS estimator structure it can be shown that the estimator will always underestimate the component. The amount of the underestimation depends on the particular model and as a rule the relative underestimation will be large when the variance of the component is relatively small, MARAVALL, A. (1995).

\(^{104}\) The theoretical variance (Estimator) should be similar to the estimate actually obtained (Estimate). Large differences between the theoretical and empirical values would indicate misspecification of the overall model, MARAVALL, A. (1995).
will be in close agreement with the theoretical ones\textsuperscript{105}. If for a given component, the variance of its theoretical estimators is significantly greater than the variance of its empirical estimate, then this component is underestimated. The opposite relationship indicates the overestimation of the component.

The outcome of the over/under estimation test is provided as a p-value in the last column of the table presented in Figure 7.46. A p-value in green denotes that the problematic characteristic has not been detected. An outcome in yellow signals the uncertain results. An outcome in red implies that an issue should be addressed. For example, when for a given component, the variance of its theoretical estimator is significantly greater than the variance of its empirical estimate, this component is underestimated. Therefore, a p-value in red manifests a strong underestimation of the component variance; a p-value in yellow is a sign of a mild underestimation of the component variance; and a p-value in green indicates no underestimation of the component variance. On the contrary, if the variance of the theoretical estimator of a component is significantly lower than the variance of its empirical estimate, this indicates that this component is underestimated. In such a case a p-value in red is interpreted as a strong overestimation of the component variance; a p-value in yellow indicates a mild overestimation of the component variance; and a p-value in green indicates no overestimation of the component variance.

**Autocorrelation functions**

The autocorrelation function (ACF)\textsuperscript{106} is a basic tool for time domain analysis of a time series. For each component and for the seasonally adjusted series, JDemetra+ exhibits autocorrelations of the stationary transformation of the components, the estimators and the sample estimates. They are calculated from the first regular lag up to the first seasonal lag. If the model is correct, the empirical estimate of the autocorrelation function should be close to the theoretical estimator of the autocorrelation function. However, for small values of innovation variance the discrepancy between the ACF function of the components and of the estimator can be substantial. In general, the more stable a component is, the larger the discrepancy\textsuperscript{107}.

For each table displayed in the Autocorrelation section the p-values of the test are given in the last column. The user should check whether the empirical estimates agree with the model, i.e. if their ACF functions are close to those of the model for the estimators. Special attention should be given to the first and/or seasonal order autocorrelation\textsuperscript{108}.

**Table 7.6: Meaning of the p-value for autocorrelation tests.**

\textsuperscript{105} MARAVALL, A. (1995).
\textsuperscript{106} See 7.9.
\textsuperscript{107} MARAVALL, A. (1993).
\textsuperscript{108} MARAVALL, A. (2000).
<table>
<thead>
<tr>
<th>Value</th>
<th>Colour</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good</td>
<td>Green</td>
<td>no evidence of autocorrelation</td>
</tr>
<tr>
<td>Uncertain</td>
<td>Orange</td>
<td>mild evidence of autocorrelation</td>
</tr>
<tr>
<td>Bad</td>
<td>Red</td>
<td>strong evidence of autocorrelation</td>
</tr>
</tbody>
</table>

It should be noted that this test does not provide information about the direction of the autocorrelation.

A comparison of the theoretical MMSE estimators with the estimates actually calculated can be used as a diagnostic tool for the model validation. The closeness of the estimators and estimates points towards validation of the results\(^\text{109}\). The failure of this test indicates the misspecification of the component models which is often due to the replacement of a non-admissible TRAMO model with its decomposable approximation, which is performed by SEATS\(^\text{110}\).

Examples of ACF are presented below.

**Trend**

<table>
<thead>
<tr>
<th>Lag</th>
<th>Component</th>
<th>Estimator</th>
<th>Estimate</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.9977</td>
<td>0.9980</td>
<td>0.9614</td>
<td>0.0000</td>
</tr>
<tr>
<td>2</td>
<td>0.9914</td>
<td>0.9921</td>
<td>0.9113</td>
<td>0.0000</td>
</tr>
<tr>
<td>3</td>
<td>0.9821</td>
<td>0.9829</td>
<td>0.8516</td>
<td>0.0000</td>
</tr>
<tr>
<td>4</td>
<td>0.9702</td>
<td>0.9710</td>
<td>0.7645</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Figure 7.47: Autocorrelations of the stationary transformation of trend, the estimators and the sample estimates.

**Seasonally adjusted**

<table>
<thead>
<tr>
<th>Lag</th>
<th>Component</th>
<th>Estimator</th>
<th>Estimate</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.9964</td>
<td>0.9972</td>
<td>0.9591</td>
<td>0.0000</td>
</tr>
<tr>
<td>2</td>
<td>0.9901</td>
<td>0.9913</td>
<td>0.9696</td>
<td>0.0000</td>
</tr>
<tr>
<td>3</td>
<td>0.9808</td>
<td>0.9820</td>
<td>0.8489</td>
<td>0.0000</td>
</tr>
<tr>
<td>4</td>
<td>0.9689</td>
<td>0.9696</td>
<td>0.7818</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Figure 7.48: Autocorrelations of the stationary transformation of the seasonally adjusted series, the estimators and the sample estimates.

**Seasonal**

<table>
<thead>
<tr>
<th>Lag</th>
<th>Component</th>
<th>Estimator</th>
<th>Estimate</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.2389</td>
<td>-0.1392</td>
<td>0.0909</td>
<td>0.8592</td>
</tr>
<tr>
<td>2</td>
<td>-0.3539</td>
<td>-0.7043</td>
<td>-0.7181</td>
<td>0.0003</td>
</tr>
<tr>
<td>3</td>
<td>-0.3849</td>
<td>-0.1566</td>
<td>-0.1595</td>
<td>0.9918</td>
</tr>
<tr>
<td>4</td>
<td>0.0000</td>
<td>0.9730</td>
<td>0.8034</td>
<td>0.0000</td>
</tr>
</tbody>
</table>


When the transitory component is not estimated, then the coefficients of the autocorrelation function of the irregular component are always zero (the Component column) as a theoretical model for irregular component \((u_t)\) is a white noise\(^{111}\). However the irregular estimator follows the ARMA model:

\[
\hat{u}_t = k_u \frac{1}{\psi(B)\psi(F)} x_t,
\]

\[\text{Var}(u)\]

where \(k_u = \text{Var}(a)\).

The estimator \(\hat{u}_t\) is expressed as a linear function of present and future innovations. Therefore, it is autocorrelated (the Estimator column contains non-zero values).

When the transitory component is estimated, the irregular and transitory component are treated together. Therefore, in such situations the coefficients of the autocorrelation function of the irregular component might be non-zero (the Component column).

### Cross-correlation function

The decomposition made by SEATS assumes orthogonal components. To test this assumption, JDemetra+ presents a table that contains cross-correlations between the stationary transformations of the components, the estimators and the actual estimates (theoretical components are uncorrelated). The cross-correlations are given for: trend and seasonal, trend and irregular, seasonal and irregular and, if the transitory is present, trend and transitory, seasonal and transitory, irregular and transitory.

\[^{111}\text{MARAVALL, A. (1987).}\]
Although components of the time series are assumed to be uncorrelated, their estimators can be correlated as the estimator variance will always underestimate the component variance. MMSE estimator implies correlation between the estimators of the components. For this reason correlations between the stationary transformations of the estimators and of the estimates actually obtained should be checked.

The first column gives the theoretical correlations. The second column provides the empirical correlation between the stationary transformations of the estimated components. Finally, the last column tests that the estimated correlations don’t differ from their theoretical value, using Bartlett’s approximations.

<table>
<thead>
<tr>
<th>Cross-correlation</th>
<th>Estimator</th>
<th>Estimate</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trend/Seasonal</td>
<td>0.0432</td>
<td>0.0281</td>
<td>0.0000</td>
</tr>
<tr>
<td>Trend/Irregular</td>
<td>0.0215</td>
<td>0.0500</td>
<td>0.0040</td>
</tr>
<tr>
<td>Seasonal/Irregular</td>
<td>0.2191</td>
<td>0.2045</td>
<td>0.7448</td>
</tr>
</tbody>
</table>

Figure 7.51: Cross correlation between the component estimators and the component sample estimates.

It is expected that the theoretical cross-correlations between the component estimators will be close to their sample estimates.

7.2.1.2.8. Significant seasonality

7.2.1.2.9. Stationary variance decomposition

---

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)

<table>
<thead>
<tr>
<th>Component</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cycle</td>
<td>2.36</td>
</tr>
<tr>
<td>Seasonal</td>
<td>29.60</td>
</tr>
<tr>
<td>Irregular</td>
<td>2.95</td>
</tr>
<tr>
<td>TD &amp; Hol.</td>
<td>1.34</td>
</tr>
<tr>
<td>Others</td>
<td>59.42</td>
</tr>
<tr>
<td>Total</td>
<td>95.06</td>
</tr>
</tbody>
</table>

Figure 7.53: Relative contribution of the components to the stationary portion of the variance in the original series with the long term trend subtracted.

7.2.1.3. Benchmarking

In the context of seasonal adjustment, benchmarking means the procedure that ensures the consistency over the calendar year between adjusted and non-seasonally adjusted data. It should be noted that the ‘ESS Guidelines on Seasonal Adjustment’ (2015) do not recommend benchmarking, as it introduces a bias in the seasonally adjusted data. Therefore, for the pre-specified specifications this option is disabled and in the results tree the Benchmarking node is empty. To activate it, click on the Specifications button, and then activate the checkbox Is enabled in the Benchmarking section and click Apply.

Figure 7.54: The Benchmarking option – a default view.

The results of the benchmarking procedure are explained in the ‘JDemetra+ User Guide’ (2016), item 3.2.1.9. The description of the benchmarking procedure is given in 7.8.
7.2.1.4. Diagnostics

The Diagnostic panel contains detailed information on the quality of the seasonal adjustment. These diagnostics are divided into five sections: Seasonality tests, Spectral analysis, Sliding spans, Revisions history and Model stability.

7.2.1.4.1. Main panel

The summary of the quality assessment is presented directly in the main panel of the Diagnostics node and is divided into several groups.

There are a range of indicators (tests) offered by JDemetra+ that measure the quality of seasonal adjustment results. This set is not identical with the sets implemented the original seasonal adjustment programs. The interpretation of the outcomes of the tests could be problematic for an inexperienced user. For this reason, the outcomes of the tests are accompanied with the values of the qualitative indicator, which facilitate the assessment of the results from the process. The values of this qualitative indicator are given in Table 7.7.

Table 7.7: The values of the qualitative indicator.

<table>
<thead>
<tr>
<th>Value</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undefined</td>
<td>The quality is undefined because of an unprocessed test, a meaningless test, a failure in the computation of the test etc.</td>
</tr>
<tr>
<td>Error</td>
<td>There is a logical error in the results (for instance, it contains aberrant values or some numerical constraints are not fulfilled). The processing should be rejected.</td>
</tr>
<tr>
<td>Severe</td>
<td>There is no logical error in the results but they should not be accepted for serious quality reasons.</td>
</tr>
<tr>
<td>Bad</td>
<td>The quality of the results is bad following a specific criterion, but there is no actual error and the results can be used.</td>
</tr>
<tr>
<td>Uncertain</td>
<td>The result of the test shows that the quality of the seasonal adjustment is uncertain.</td>
</tr>
<tr>
<td>Good</td>
<td>The result of the test is good from the aspect of the quality of seasonal adjustment.</td>
</tr>
</tbody>
</table>

Several qualitative indicators can be combined following the basic, arbitrary rules. In particular, it is done for the Summary indicator, which gives the first insight into the quality of the estimation and can be especially useful for the assessment of the seasonal adjustment of large datasets performed under time pressure. The quality assessment provided by the Summary indicator aims to support the user when the high number of series that are being treated makes it hard or even
impossible to perform an individual analysis of the results, although an additional validation of the results by the user is recommended.

Figure 7.55: The value of the summary indicator for a given series.

The rule for the calculation of the Summary indicator as well as other aggregated indicators, which combine $n$ qualitative indicators, is given in Table 7.8. To calculate the average of the (defined) diagnostics, 0 is assigned to Bad, 2 is assigned to Uncertain and 3 is assigned to Good.

Table 7.8: The overall value of the aggregated indicator.

<table>
<thead>
<tr>
<th>Sum</th>
<th>Rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undefined</td>
<td>All of $n$ qualitative indicators are Undefined.</td>
</tr>
<tr>
<td>Error</td>
<td>The value of at least one of $n$ qualitative indicators is Error.</td>
</tr>
<tr>
<td>Severe</td>
<td>None of the $n$ qualitative indicators is Error and at least one of them is Severe.</td>
</tr>
<tr>
<td>Bad</td>
<td>None of the $n$ qualitative indicators is Error or Severe. The average of the (defined) diagnostics is less than 1.5.</td>
</tr>
<tr>
<td>Uncertain</td>
<td>None of the $n$ qualitative indicators is Error or Severe. The average of the (defined) diagnostics is in the range [1.5, 2.5].</td>
</tr>
<tr>
<td>Good</td>
<td>None of the $n$ qualitative indicators is Error or Severe. The average of the (defined) diagnostics is at least 2.5.</td>
</tr>
</tbody>
</table>

According to Table 7.8, the Error and Severe diagnostics are unacceptable results.

Under the Summary indicator a set of tests that check for the presence of residual seasonality is displayed. The residual seasonality diagnostics implemented in JDemetra+ correspond to the set of tests developed for X-12-ARIMA. One of them is the F-test for the presence of residual seasonality, which is computed on the detrended seasonally adjusted series and on the irregular component.

\begin{align*}
\text{residual seasonality} \\
\text{on sa: Good (0.994)} \\
\text{on sa (last 3 years): Good (0.890)} \\
\text{on irregular: Good (0.846)}
\end{align*}

Figure 7.56: Residual seasonality test results.

In order to remove the trend from a monthly time series, a first order difference of lag three is applied (a first order difference of lag one in the other cases)\(^{116}\) to the seasonally adjusted series.

For the detrended seasonally adjusted series the presence of residual seasonality is tested on the complete time span and on the last 3 years span. The description of the F-test is given in 7.6.2.1.

Table 7.9: The threshold values for the results of the F-test for the presence of residual seasonality.

<table>
<thead>
<tr>
<th>P-value</th>
<th>JDemetra+ default setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;0.01</td>
<td>Severe</td>
</tr>
<tr>
<td>[0.01, 0.05]</td>
<td>Good</td>
</tr>
<tr>
<td>[0.05, 0.1]</td>
<td>Uncertain</td>
</tr>
<tr>
<td>≥0.1</td>
<td>Good</td>
</tr>
</tbody>
</table>

Indicators in the seats section check if the crucial assumptions concerning the relationship between the components are fulfilled. These tests are discussed in 5.2.1.2.7.

Figure 7.57: The summary results from decomposition performed by SEATS.

The out-of-sample section summarises the results for the tests performed on forecasts, discussed in 4.2.3.

Figure 7.58: The summary results from out-of-sample tests.

Next, several tests are computed on the residuals of the TRAMO model. The definition of the residuals slightly differs from those of the original X-13ARIMA-SEATS and the TRAMO/SEATS algorithms. However, their global messages are nearly always very similar.

Figure 7.59: Tests on residuals from the RegARIMA model.

The normality test (which combines skewness and kurtosis tests), displayed in Figure 7.59, is the Doornik-Hansen test (see 7.6.1.1), which follows a \( \chi^2 \) distribution.

Table 7.10: The threshold values for the results of the Doornik-Hansen normality test.
The independence test is the Ljung-Box test (see 7.6.1.3), which follows a $\chi^2_{(k-np)}$ distribution, where $k$ depends on the frequency of the series (24 for a monthly series, 8 for a quarterly series, $4 \times freq$ for other frequencies, where $freq$ is a frequency of the time series) and $np$ is the number of hyper-parameters of the model (number of parameters in the ARIMA model).

<table>
<thead>
<tr>
<th>Pr($\chi^2 &gt; val$)</th>
<th>JDemtra+ default setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 0.01</td>
<td>Bad</td>
</tr>
<tr>
<td>[0.01, 0.1]</td>
<td>Uncertain</td>
</tr>
<tr>
<td>$\geq 0.1$</td>
<td>Good</td>
</tr>
</tbody>
</table>

Table 7.11: The threshold values for the results of the Ljung-Box test.

JDemtra+ checks for the presence of trading day (spectral td peaks) and seasonal peaks (spectral seas peaks) in the residuals using a test based on the periodogram of the residuals. The periodogram is computed at the so-called Fourier frequencies. Under the hypothesis of Gaussian white noise of the residuals, it is possible to derive a simple test on the periodogram, around specific (groups of) frequencies. The exact definition of the test is described in 7.6.2.9.

<table>
<thead>
<tr>
<th>P(\text{stat} &gt; val)</th>
<th>JDemtra+ default setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 0.001</td>
<td>Severe</td>
</tr>
<tr>
<td>[0.001, 0.01]</td>
<td>Bad</td>
</tr>
<tr>
<td>[0.01, 0.1]</td>
<td>Uncertain</td>
</tr>
<tr>
<td>$\geq 0.1$</td>
<td>Good</td>
</tr>
</tbody>
</table>

Table 7.12: The threshold values for the results of the test on periodogram.

JDemtra+ identifies spectral peaks in the seasonal and the trading day components using an empirical criterion of "visual significance". For more information see 7.3.2.

```
visual spectral analysis
spectral seas peaks: Good (0.000)
spectral td peaks: Good (0.000)
```

Figure 7.60: The summary results from visual spectral analysis tests.

The **basic checks** section includes two quality diagnostics: definition and annual totals.
The definition test inspects some basic relationships between different components of the time series. In the case of an additive decomposition the following relationships are checked:

- \( mhe = ee + omhe \) \ [5.8]\n- \( cal = tde + mhe \) \ [5.9]\n- \( out = out_t + out_s + out_i \) \ [5.10]\n- \( reg = reg_t + reg_s + reg_i + reg_y \) \ [5.11]\n- \( reg.sa = reg_t + reg_i \) \ [5.12]\n- \( det = cal + out + reg \) \ [5.13]\n- \( c.t = t + o.t + reg.t \) \ [5.14]\n- \( c.s = s + cal + o.s + reg.s \) \ [5.15]\n- \( c.i = i + o.i + reg.i \) \ [5.16]\n- \( c.sa = y.c - c.s = c.t + c.i + reg.y \) \ [5.17]\n- \( cy.c = c.t + c.s + reg.y = t + s + i + reg \) \ [5.18]\n- \( y.l = y.c - det = t + s + i \) \ [5.19]\n- \( sa = y - s = t + i \) \ [5.20]\n- \( si = y.l - t = s + i \) \ [5.21]\n
The explanations of the abbreviation used in these formulas are given in Table 7.13.
Table 7.13: The components and effects used in the definition test formulas.

<table>
<thead>
<tr>
<th>Name</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>( y )</td>
<td>Original series</td>
</tr>
<tr>
<td>( y_c )</td>
<td>Interpolated time series (i.e. original time series with missing values replaced by their estimates)</td>
</tr>
<tr>
<td>( t )</td>
<td>Trend (without regression effect)</td>
</tr>
<tr>
<td>( s )</td>
<td>Seasonal (without regression effect)</td>
</tr>
<tr>
<td>( i )</td>
<td>Irregular (without regression effect)</td>
</tr>
<tr>
<td>( sa )</td>
<td>Seasonally adjusted series (without regression effect)</td>
</tr>
<tr>
<td>( si )</td>
<td>S-I ratio</td>
</tr>
<tr>
<td>( tde )</td>
<td>Trading day (or working day) effect</td>
</tr>
<tr>
<td>( mhe )</td>
<td>Moving holidays effect</td>
</tr>
<tr>
<td>( ee )</td>
<td>The Easter effect</td>
</tr>
<tr>
<td>( omhe )</td>
<td>Other moving holidays effect</td>
</tr>
<tr>
<td>( cal )</td>
<td>Calendar effects</td>
</tr>
<tr>
<td>( out )</td>
<td>Total outlier effect</td>
</tr>
<tr>
<td>( out_t )</td>
<td>Effect of outliers assigned to the trend component (LS)</td>
</tr>
<tr>
<td>( out_s )</td>
<td>Effect of outliers assigned to the seasonal component (SO)</td>
</tr>
<tr>
<td>( out_i )</td>
<td>Effect of outliers assigned to the irregular component (AO, TC)</td>
</tr>
<tr>
<td>( reg )</td>
<td>Effect of regression variables (except for outliers)</td>
</tr>
<tr>
<td>( reg_t )</td>
<td>Effect of regression variables (except for outliers) assigned to the trend component</td>
</tr>
<tr>
<td>( reg_s )</td>
<td>Effect of regression variables (except for outliers) assigned to the seasonal component</td>
</tr>
<tr>
<td>( reg_i )</td>
<td>Effect of regression variables (except for outliers) assigned to the irregular component</td>
</tr>
<tr>
<td>( reg_y )</td>
<td>Separate regression effects (except for outliers)</td>
</tr>
<tr>
<td>( reg_sa )</td>
<td>Effect of regression variables (except for outliers) assigned to the seasonally adjusted series</td>
</tr>
<tr>
<td>( det )</td>
<td>Deterministic effect</td>
</tr>
<tr>
<td>( c_t )</td>
<td>Trend, including deterministic effect</td>
</tr>
<tr>
<td>( c_i )</td>
<td>The irregular component, including deterministic effect</td>
</tr>
<tr>
<td>( c_s )</td>
<td>The seasonal component, including deterministic effect</td>
</tr>
<tr>
<td>( c_y )</td>
<td>The original series, including deterministic effect</td>
</tr>
<tr>
<td>( c_sa )</td>
<td>The seasonally adjusted series, including deterministic effect</td>
</tr>
<tr>
<td>( cal_y )</td>
<td>Calendar adjusted series</td>
</tr>
</tbody>
</table>
A multiplicative model \(^{117}\) is obtained in the same way by replacing the operations "+" and "-" by "*" and "/" respectively. The explanations of the abbreviations are given in 7.7.

The definition test verifies that all the definition constraints are well respected. The maximum of the absolute differences is computed for the different equations and related to the Euclidean norm of the initial series (\(Q\)).

<table>
<thead>
<tr>
<th>(Q)</th>
<th>Diagnostic</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; 0.000001</td>
<td>Error</td>
</tr>
<tr>
<td>&lt;= 0.000001</td>
<td>Good</td>
</tr>
</tbody>
</table>

The annual totals test compares the annual totals of the original series with those of the seasonally adjusted series. The maximum of their absolute differences is computed and related to the Euclidean norm of the initial series.

<table>
<thead>
<tr>
<th>(Q)</th>
<th>Diagnostic</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; 0.5</td>
<td>Error</td>
</tr>
<tr>
<td>[0.1, 0.5]</td>
<td>Severe</td>
</tr>
<tr>
<td>[0.05, 0.1]</td>
<td>Bad</td>
</tr>
<tr>
<td>[0.01, 0.05]</td>
<td>Uncertain</td>
</tr>
<tr>
<td>&lt;=0.01</td>
<td>Good</td>
</tr>
</tbody>
</table>

7.2.1.4.2. Seasonality tests

The Diagnostics node includes a set of seasonality tests that are useful for checking for the presence of seasonality in time series. The tests are described in 7.6.2 and in the ‘JDemetra+ User Guide’ (2016), item 3.4. The tests check for the presence of seasonality in:

- The original series (log transformed, if necessary);
- Linearised series;
- Full residuals;

\(^{117}\) See 7.1.1.
- SA series;
- Irregular;
- Residuals (last periods);
- SA series (last periods);
- Irregular (last periods);

Seasonal adjustment should be applied only to seasonal series. However, regression effects can mask the underlying seasonal pattern. Therefore, seasonality should be statistically significant in the linearised series. Finally, seasonally adjusted series should have neither residual seasonality nor residual calendar effects and should show both the full trend-cycle and irregular component.\textsuperscript{118} The presence of seasonality can also be checked by the Combined seasonality test, which originates from X-11 algorithm and considers the results of several seasonality tests to judge if the series includes identifiable seasonal fluctuations.

7.2.1.4.3. Sliding spans

Seasonally adjusted time series are expected to be stable, which means that they do not change substantially when removing or adding few data points to the original series. This characteristic is assessed by the sliding spans\textsuperscript{127} analysis. This tool can be also used to detect significant changes in the original time series, such as seasonal breaks\textsuperscript{119}, large number of outliers and fast moving seasonality\textsuperscript{120}.

A sliding spans analysis uses the concept of a span, which is defined as a range of data between two dates. The sliding spans are series of two, three or four, depending on the length of the original time series overlapping spans. The program sets up a maximum of 4 spans, the length of each span is always 8 years\textsuperscript{121}. The spans start in 1 year intervals.

\textsuperscript{118} ‘ESS Guidelines on Seasonal Adjustment’ (2015).\textsuperscript{127}
See 7.4.

\textsuperscript{119} A seasonal break is defined as a sudden and sustained change in the seasonal pattern of a series. The presence of this event is reflected in SI ratio. A seasonal breaks are unwanted feature of the time series as the moving averages used by X-1-ARIMA are designed to deal with series which have a smoothly evolving ‘deterministic’ seasonal component plus an irregular component with stable variance.

\textsuperscript{120} Fast moving seasonality means that the seasonal pattern displays rapidly evolving fashion from year to year.

\textsuperscript{121} The procedure of withdrawing spans from time series is described in FINDLEY, D., MONSELL, B.C., SHULMAN, H.B., and PUGH, M.G. (1990) as follows: to obtain sliding spans for a given series, an initial span is selected whose length depends on the seasonal adjustment filters being used. A second span is obtained from this one by deleting the earliest year of data and appending the year of data following the last year in the span. A third span is obtained from the second in this manner, and a fourth from the third, data permitting. This is done in such a way that the last span contains the most recent data.
The sliding spans analysis compares the seasonally adjusted values for a given observation, obtained by applying the seasonal adjustment procedure to the consecutive spans. Each period (month or quarter), that is common to more than one span, is examined to see if its seasonal adjustments vary more than a specified amount across the spans.

The summary of the sliding spans includes the results of several seasonality tests for each span. The differences between spans signal the change in the characteristic of seasonal movements in the time series. For the description of the seasonality tests see 7.4.

### Tests for seasonality

<table>
<thead>
<tr>
<th></th>
<th>Span 1</th>
<th>Span 2</th>
<th>Span 3</th>
<th>Span 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stable seas.</td>
<td>356.6</td>
<td>258.4</td>
<td>316.9</td>
<td>194.1</td>
</tr>
<tr>
<td>Kruskal-Wallis</td>
<td>94.4</td>
<td>93.1</td>
<td>94.1</td>
<td>91.9</td>
</tr>
<tr>
<td>Moving seas.</td>
<td>4.1</td>
<td>2.2</td>
<td>1.2</td>
<td>2.1</td>
</tr>
<tr>
<td>Identifiable seas.</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

The differences in the mean of seasonal factors between different spans observed for a given period signal a potential change in the characteristic of the seasonal fluctuations.

### Means of seasonal factors

<table>
<thead>
<tr>
<th>Month</th>
<th>Span 1</th>
<th>Span 2</th>
<th>Span 3</th>
<th>Span 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>0.5985</td>
<td>0.6017</td>
<td>0.6029</td>
<td>0.6085</td>
</tr>
<tr>
<td>February</td>
<td>0.8010</td>
<td>0.8001</td>
<td>0.8089</td>
<td>0.8233</td>
</tr>
<tr>
<td>March</td>
<td>1.3413</td>
<td>1.3334</td>
<td>1.3446</td>
<td>1.3783</td>
</tr>
<tr>
<td>April</td>
<td>1.5288</td>
<td>1.4752</td>
<td>1.4687</td>
<td>1.4666</td>
</tr>
<tr>
<td>May</td>
<td>1.3612</td>
<td>1.3511</td>
<td>1.3537</td>
<td>1.3296</td>
</tr>
<tr>
<td>June</td>
<td>1.2653</td>
<td>1.2506</td>
<td>1.2341</td>
<td>1.2223</td>
</tr>
<tr>
<td>July</td>
<td>1.1537</td>
<td>1.1607</td>
<td>1.1655</td>
<td>1.1638</td>
</tr>
<tr>
<td>August</td>
<td>0.9969</td>
<td>1.0120</td>
<td>1.0115</td>
<td>0.9982</td>
</tr>
<tr>
<td>September</td>
<td>0.8903</td>
<td>0.9025</td>
<td>0.9013</td>
<td>0.9079</td>
</tr>
<tr>
<td>October</td>
<td>0.7742</td>
<td>0.8000</td>
<td>0.8172</td>
<td>0.8220</td>
</tr>
<tr>
<td>November</td>
<td>0.6844</td>
<td>0.7050</td>
<td>0.7012</td>
<td>0.6883</td>
</tr>
<tr>
<td>December</td>
<td>0.6061</td>
<td>0.6002</td>
<td>0.5940</td>
<td>0.5992</td>
</tr>
</tbody>
</table>
JDemetra+ performs sliding spans analysis not only for seasonally adjusted series, but also for the trading day/working day effect (if present) and the seasonal component. The detailed results are displayed in the sub-nodes: Seasonal, Trading days and SA (changes). The later panel is related to the period-to-period percentage changes in seasonally adjusted series. The threshold value to detect abnormal values is set to 3% of the testing statistics. The respective formulas are given in 7.4. In this section the standard Copy/Print/Export options are available only for graph that presents the sliding spans statistics.

The layout of each of these sub-nodes is the same. The explanations provided here for the seasonal component can be adapted accordingly to the other two sub-nodes. The user should be aware that an unstable estimate of a seasonal factor for a given period can give rise to unstable estimates of the two associated period-to-period changes. Because of that, in the majority of cases, more months are flagged for unreliable period-to-period changes than for unreliable seasonal factors122.

The first panel shows the sliding spans statistics calculated for each period. For the given period the sliding spans statistic is the maximum percentage difference between the estimates of the seasonal component obtained from different spans. The estimation of seasonal component is regarded as unstable if this statistic is greater than 3%.

![Figure 7.65: The sliding spans statistics for the seasonal component.](image)

The next panel presents the cumulative frequency distribution of the sliding spans statistics (months or quarters) using the frequency polygon. On the horizontal axis the values of the sliding spans statistics are shown, while vertical axis presents the frequency (in percentages) of each class interval123. Figure 7.66 shows distribution where the first interval extends from 0 to 0.005. This interval has a frequency 9%, which means that 9% of the sliding spans statistics are in this interval.


123 In frequency polygon data presented on the horizontal axis are grouped into class intervals.
According to FINDLEY, D., MONSELL, B.C., SHULMAN, H.B., and PUGH, M.G. (1990), the results of seasonal adjustment are stable if the percentage of unstable (abnormal) seasonal factors is less than 15% of the total number of observations. Empirical surveys support the view that seasonal adjustments with more than 25% of the months (or quarters) flagged for unstable seasonal factor estimates are not acceptable. Therefore, the user should check for the total frequency of the intervals between 0.03 and 1.

The last panel contains detailed information about the percentage of values that exceed the threshold value (0.03) and therefore are considered to be abnormal. In the example presented in Figure 7.67, 28.6% of values have been marked by the sliding spans diagnostic as abnormal.

JDemetra+ also provides information about the number of breakdowns of unstable factors and average maximum percentage differences grouped by month (or quarter) and by year. It gives an idea of whether observations with unreliable seasonal adjustment values cluster in certain calendar periods and whether their sliding spans statistics barely or substantially exceed the threshold value. Figure 7.67 presents that three sliding spans statistic calculated for January have been above 3% and that the average maximum percentage difference across spans for this period was 2.0.

---

Figure 7.67: Breakdowns of unstable factors and average maximum percent differences across spans.

Sliding spans analysis is a tool that can be used for correcting the model. For example, a large number of unstable estimates revealed by the sliding spans analysis indicates that the model specification may need to be changed. Comparing the seasonality tests results for consecutive spans may also indicate that the seasonal pattern has changed substantially. In such a case the user should investigate this issue further and decide whether to reduce the time series span.

<table>
<thead>
<tr>
<th>Year</th>
<th>Breakdowns</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>2</td>
<td>1.7</td>
</tr>
<tr>
<td>2006</td>
<td>6</td>
<td>2.5</td>
</tr>
<tr>
<td>2007</td>
<td>4</td>
<td>2.9</td>
</tr>
<tr>
<td>2008</td>
<td>3</td>
<td>2.5</td>
</tr>
<tr>
<td>2009</td>
<td>3</td>
<td>2.3</td>
</tr>
<tr>
<td>2010</td>
<td>3</td>
<td>2.3</td>
</tr>
<tr>
<td>2011</td>
<td>6</td>
<td>3.0</td>
</tr>
<tr>
<td>2012</td>
<td>1</td>
<td>1.7</td>
</tr>
<tr>
<td>2013</td>
<td>3</td>
<td>1.6</td>
</tr>
<tr>
<td>2014</td>
<td>1</td>
<td>2.8</td>
</tr>
</tbody>
</table>

Figure 7.68: Sliding span summary for the series for which the seasonal fluctuations fade away.

7.2.1.4.4. Spectral analysis

Seasonal and calendar effects are approximately periodic, therefore the spectrum is an appropriate
their presence\textsuperscript{125}. The periodicity of phenomenon at frequency $f$ is $\frac{2\pi}{f}$. It means that $\frac{\pi}{2}, \frac{3\pi}{2}, \frac{4\pi}{2}, \frac{5\pi}{2}$ for a monthly time series the seasonal frequencies are: $\frac{\pi}{6}, \frac{2\pi}{6}, \frac{3\pi}{6}, \frac{4\pi}{6}, \frac{5\pi}{6}$ (which are equivalent to $\frac{\pi}{6}, \frac{2\pi}{6}, \frac{3\pi}{6}, \frac{4\pi}{6}, \frac{5\pi}{6}$). For a quarterly series there are two seasonal frequencies: $\frac{\pi}{2}$ (one cycle per year) and $\frac{\pi}{2}$ (two cycles per year). A peak at the zero frequency always corresponds to the trend component of the series. For more details about spectral analysis see 7.3.

JDemetra+ displays the spectral plots obtained for two spectrum estimators: periodogram and autoregressive spectrum\textsuperscript{126}. They are intended to alert to the presence of remaining seasonal and trading day effects. The graphics are available for the residuals, the irregular component and the seasonally adjusted time series. Seasonal frequencies are marked as blue vertical lines, while red lines represent the trading day effect frequencies. The $X$-axis shows the frequencies from 0 to $\pi$.

The interpretation of the spectral graph is rather straightforward. When the values of a spectral graph for low frequencies are large in relation to the other values it means that the long-term movements dominate in the series. When the values of a spectral graph for high frequencies are large in relation to its other values it means that the series is rather trendless and contains a lot of noise. When the values of a spectral graph are distributed randomly around a constant without any visible peaks, then it is highly probable that the series is a random process. The presence of seasonality in a time series is manifested by the peaks at the seasonal frequencies while the peaks at the trading day frequencies correspond to the presence of a trading day effect. The explanations for the trading day frequencies are given in 7.6.2.9.

At the seasonal and trading day frequencies, a peak in the residuals indicates the need for fitting a better model. In particular, peaks at the seasonal frequencies are caused by inadequate filters chosen for the decomposition. Peaks at the trading day frequencies could occur due to inappropriate regression variables used in the model.

A peak in the spectrum of the seasonally adjusted series or irregulars reveals inadequacy of the seasonal adjustment filters for the time interval used for spectrum estimation. In this case a different model specification or data span length should be considered.

\textsuperscript{125} FINDLEY, D., MONSELL, B.C., BELL, W.R., OTTO, M.C., and CHEN, B.-C. (1990). Basic information about spectral analysis is presented in 7.3.

\textsuperscript{126} The theoretical motivation for the choice of spectral estimator is provided by SOKUP, R., and FINDLEY, D. (1999).
If at least one significant peak ($P_{value} > 0.05$) has been detected in the residuals, JDemetra+ displays a green line that denotes the 0.05 significance level. Peaks above this limit are considered to be significant at the 0.05 significance level\textsuperscript{127}.

An autoregressive spectrum graph available in JDemetra+ is based on the relevant tool from the X13ARIMA-SEATS program. It shows the spectral density (spectrum) function which reformulates the autocovariances of the stationary time series in terms of amplitudes at frequencies of half a cycle per month or less.

![Autoregressive spectrum](image)

Figure 7.69: Autoregressive spectrum for residuals from the TRAMO model.

The periodogram is one of the earliest tools used for the analysis of time series in the frequency domain. It enables us to identify the dominant periods (or frequencies) of a time series. In general the periodogram is a wildly fluctuating estimate of the spectrum with high variance and is less stable than the autoregressive spectrum\textsuperscript{128}.

![Periodogram](image)

Figure 7.70: Periodogram for residuals from TRAMO model.

For computational reasons none of the periodograms for seasonally adjusted series and for the irregular displays the line that indicates the peaks above the 0.05 significance level. Therefore, there is no message if the peaks visible on those panels are significant or not.

Standard Copy/Print/Export options are available for these charts.

\textsuperscript{127} This indicator appears only if the series is a white noise (residuals) as the periodogram is then distributed as a $\chi^2$.

\textsuperscript{128} For description of the periodogram see 7.3.1. The autoregressive spectrum is discussed in 7.3.2.
7.2.1.4.5. Revision histories

It is well-known that the seasonally adjusted and trend estimates change as new observations are added to the end of the original time series. The changes in the estimated SA and trend values are called revisions and can be analysed with the tool called Revision history. JDemtra+ illustrates differences between the initial estimate (marked by a blue circle) and the latest estimate (red line). A revision is defined as a difference between those two values. As a rule, smaller revisions are better.

The revision history is useful for comparing results from competing models. When the user defines two seasonal adjustment models for one time series and both these models are acceptable, the revision history can be used for choosing a better model in terms of revisions. However, it is not an actual statistical test but a complementary descriptive analysis.

The revision history is performed for the maximum of 84 most recent observations (monthly series) or 16 most recent observations (quarterly series). When a time series is shorter than 109 observations (monthly series) or 37 observations (quarterly series) the number of data points considered by the revision history analysis is reduced accordingly. For series shorter than 62 observations (monthly series) or 22 observations (quarterly series) the revision history is not computed.

If the user clicks on a blue circle, which represents the initial estimation for period \( t_n \), an auxiliary window will appear. This additional figure shows the successive estimations (computed on \([t_0, \ldots, t_n], [t_0, \ldots, t_{n+1}], \ldots , [t_0, \ldots, t_T]\) of the considered series for the period \( t_n \). With this figure the user can evaluate how the seasonally adjusted observations have changed from the initial to the final estimation. The analogous graph is available for the trend analysis.
By looking at the vertical axis the user could judge the size of the revision. The revisions presented in Figure 7.72 are about 2% (103 - 101 = 2). The figure size can be enlarged by dragging the bottomright corner.

The revision history analysis plot is accompanied by information about the relative differences between the initial and the final estimates for the last four years. For an additive decomposition the absolute revisions are used. For a multiplicative decomposition the relative differences are considered. Values, which are larger (in absolute terms) than two times the root mean squared error of the (absolute or relative) revisions, are marked in red and provide information about the instability of the output. Information about the mean relative difference between the initial and the final estimation over a period displayed in the table is also provided. As relative differences can be positive as well as negative, the mean value is not very informative for detecting any possible bias. The magnitude of varying revisions is measured by the root mean square error (RMSE), which has the same units as the mean.
The revision history provides no absolute measure of what is an acceptable level of revisions. Therefore the revision history is of limited use on a single series. More information about the revision history is available in 7.5.

7.2.1.4.6. Model stability

The diagnostics output window provides some purely descriptive features which can be used to analyse the stability of the parameters of the TRAMO model, including the trading/working day effect, the Easter and the ARIMA part. The model stability analysis estimates the coefficients of the TRAMO model applied for 8 year long spans. The number of estimations depends on the length of the series. For example, when the series is 12 years long, then only five estimations will be calculated. The points displayed in the figure correspond to the successive estimations.

The figures displayed in this section are useful for assessing the stability of the model parameters. The first panel presents estimation results for one (working day effect) or six (trading day effect) regression coefficients. The graph is not displayed when none of these effects is included in the TRAMO model.
Figure 7.74: Estimation parameters for the trading day effect.

The concentration of the estimation results (denoted by blue circles) around the mean (red, horizontal line) indicates the stability of the results in time. In the case presented in Figure 7.75 however, the values of the Easter effect parameters vary from 2.5 to -3.5, indicating the instability of the results and potential insignificance of this effect.

Figure 7.75: Estimation parameters for the Easter effect.

The results from the estimation of an ARIMA (3,1,0)(0,0,1) model over the 8 spans, which is shown in Figure 7.76, reveal quite high stability of the estimations.
The user can display separately the results for a given parameter. It can be done by double clicking on the area of interest. The details are displayed in a resizable pop-up window (double right click to the restore the default view).

7.2.2. X13

The X-13ARIMA-SEATS method consists of two linked parts: the RegARIMA model and the decomposition step that can be performed using the X-11 algorithm or the SEATS program. The results from RegARIMA, which are displayed under the Pre-processing node, are explained in 4.2. The output from the decomposition step is presented in the three nodes: Decomposition, Benchmarking and Diagnostics. The majority of indicators displayed in the Diagnostics is shared with TRAMO/SEATS and are discussed in 5.2.1.4. The Benchmarking section is covered in 5.2.1.3. For the Main results node only one section is different from the output produced by TRAMO/SEATS. This section focuses on the nodes that are not handled elsewhere, i.e. the output produced by X-11. Therefore the sections that are explained here are Decomposition and some parts of the Main results and Diagnostics sections. Their content is accessible once the user selects the appropriate node from the left-hand side seasonal adjustment results panel.
7.2.2.1. Main results

The first section of the Main results node summarises the results of the pre-processing. The content of this panel depends on the specifications used for processing and the results of the seasonal adjustment\textsuperscript{129}. In the case of the pre-defined specifications X11, RSA0, RSA1 and RSA3 no trading day effect is estimated. For RSA2c and RSA4c pre-defined specifications, working day effects and the leap year effect are pre-tested and estimated if present. If the working day effect is significant, the pre-processing part includes the message Working days effects (1 variable). The message Working days effects (2 variables) means that the leap year effect has also been estimated. For RSA5c the trading day effect and the leap year effect are pre-tested. If the trading day effect has been detected, either of the messages Trading days effects (6 variables) or Trading days effects (7 variables) are displayed, depending whether the leap year effect has been detected or not. If the Easter effect is statistically significant, Easter effect detected is displayed. In this section, only the total number of detected outliers is visible. More detailed information, i.e. type, location and coefficients of every outlier, can be found in the Pre-processing node.

\textsuperscript{129} For description of the pre-defined specifications see 4.1 and 5.1. User-defined specifications are described in ‘JDemetra+ User Guide’ (2016), item 3.2.1.
Pre-processing [RegARima]

Series has been log-transformed
No trading days effects
Easter effect detected
3 outliers detected

Figure 7.78: The summary of the results from RegARIMA.

The message *Series has been log-transformed* is only displayed if a logarithmic transformation has been applied. Otherwise, the message does not appear.

The $m$-statistics section provides summary indicators ($q$ and $q - m2$) from analysis of M-statistics. The details of these measures are given in 5.2.2.1.5. The result displayed in green indicates that the given indicator’s value has been accepted. When the indicator is above one the test fails and the statistic is displayed in red.

\[
\begin{align*}
\text{m-statistics} \\
q & : \text{Good (0.622)} \\
q-m2 & : \text{Good (0.703)}
\end{align*}
\]

Figure 7.79: The summary of m-statistics.

7.2.2.2. Decomposition (X11)

This part includes tables with results from consecutive iterations of the X-11 algorithm and the set of quality measures.

7.2.2.2.1. Tables

The decomposition step of X-13ARIMA-SEATS is performed by an iterative algorithm X-11, which general principle is an estimation of the different components using appropriate moving averages. The results of each step of the algorithm are saved in the successive tables, which are displayed in this section. However, some tables produced by the original X-11 algorithm are omitted. The tables are divided into six groups, which correspond to the main steps of the algorithm:

- Part A: Pre-adjustments;
- Part B: First automatic correction of the series;
- Part C: Second automatic correction of the series;
- Part D: Seasonal adjustment;
- Part E: Components modified for large extreme values;
- Part F: Seasonal adjustment quality measures.

As an example the view of the D-table group is presented below.
By default, each table is displayed as a time series. To switch into table view select the \textit{Single time series} option from the local menu.

Then, to display a given table expand the menu, which is available above the table, and choose table’s name.
A full list of the X-11 tables displayed by JDemetra+ can be found in 7.1.3.2.4. The description of the all options from the local menu is available below Figure 4.11.

### 7.2.2.2.2. Final filters

The length of the seasonal and trend moving average filters used to estimate the final seasonal factors and the final trend depends on the time series and is selected automatically by the X-11 algorithm on a basis of the time series properties. In short, JDemetra+ selects the filters automatically, taking into account the global moving seasonality ratio, which is computed on preliminary estimates of the irregular component and of the seasonal. The selection criteria are discussed in 7.1.1.3. In the case of user-defined seasonal adjustment specifications, the lengths of the seasonal and trend filters can be chosen manually (see 5.1.2.1).

#### Final filters

| Trend filter: 13-term Henderson moving average |
| Seasonal filter: 3 x 5 moving average |

Figure 7.83: Information about filters used by the X-11 algorithm to calculate the final estimate of the seasonal component and trend.

For user-defined specifications the different seasonal filters can be applied for each period (see 5.1.2.1). In such a case JDemetra+ informs that the composite filter has been applied. The orders of seasonal filters can be checked in the Specification panel in the X11 → Details on seasonal filter section.

Figure 7.84: Information about filters used by the X-11 algorithm to calculate the final estimate of the seasonal component and trend – the case of composite filters.
7.2.2.2.3. Quality measures

The quality measures are tools that can be used to assess the quality of the decomposition and determine the steps for improvement, if necessary. The set of these measures includes 11 so called $M$ statistics and two synthetic measures: $Q$ and $Q - M2$.

The $M$ statistics are used to assess the quality of the seasonal adjustment\(^\text{130}\). These statistics vary between 0 and 3 but only values smaller than 1 are acceptable. JDemetra+ displays their results together with the composite indicators $Q$ and $Q - M2$. Results displayed in red indicate the failure of a test.

<table>
<thead>
<tr>
<th>$M$</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M1$</td>
<td>The relative contribution of the irregular over three months span</td>
</tr>
<tr>
<td>$M2$</td>
<td>The relative contribution of the irregular component to the stationary portion of the variance</td>
</tr>
<tr>
<td>$M3$</td>
<td>The amount of period to period change in the irregular component as compared to the amount of period to period change in the trend-cycle</td>
</tr>
<tr>
<td>$M4$</td>
<td>The amount of autocorrelation in the irregular as described by the average duration of run</td>
</tr>
<tr>
<td>$M5$</td>
<td>The number of periods it takes the change in the trend cycle to surpass the amount of change in the irregular</td>
</tr>
<tr>
<td>$M6$</td>
<td>The amount of year to year change in the irregular as compared to the amount of year to year change in the seasonal</td>
</tr>
<tr>
<td>$M7$</td>
<td>The amount of moving seasonally present relative to the amount of stable seasonality</td>
</tr>
<tr>
<td>$M8$</td>
<td>The size of the fluctuations in the seasonal component throughout the whole series</td>
</tr>
<tr>
<td>$M9$</td>
<td>The average linear movement in the seasonal component throughout the whole series</td>
</tr>
<tr>
<td>$M10$</td>
<td>The size of the fluctuations in the seasonal component in the recent years</td>
</tr>
<tr>
<td>$M11$</td>
<td>The average linear movement in the seasonal component in the recent years</td>
</tr>
<tr>
<td>$Q$</td>
<td>0.863</td>
</tr>
<tr>
<td>$Q-M2$</td>
<td>0.967</td>
</tr>
</tbody>
</table>

Figure 7.85: X-11 quality measures.

The details about the measures are given below.

- **$M1$** measures the contribution of the irregular component to the total variance. When it is above 1 some changes in outlier correction should be considered.
- **$M2$**, which is very similar to $M1$, is calculated on the basis of the contribution of the irregular component to the stationary portion of the variance. When it is above 1 some changes in outlier correction should be considered.
- **$M3$** compares the irregular to the trend taken from a preliminary estimate of the seasonally adjusted series. If this ratio is too large, it is difficult to separate the two components from each other. When it is above 1 some changes in outlier correction should be considered.
- **$M4$** tests the randomness of the irregular component. A value above 1 denotes correlation in the irregular component. In such a case a shorter seasonal moving average filter should be considered.
- **$M5$** is used to compare the significance of changes in the trend with that in the irregular. When it is above 1 some changes in outlier correction should be considered.
- **$M6$** checks the $SI$ (seasonal - irregular components ratio). If annual changes in the irregular component are too small in relation to the annual changes in the seasonal component, the

\(^\text{130}\) For the definitions of the $M$ and $Q$ statistics see LADIRAY, D., and QUENNEVILLE, B. (1999).
3 × 5 seasonal filter used for the estimation of the seasonal component is not flexible enough to follow the seasonal movement. In such a case a longer seasonal moving average filter should be considered. It should be underlined that \( M6 \) is calculated only if this filter has been applied in the model.

- \( M7 \) is the combined test for the presence of identifiable seasonality. The test compares the relative contribution of stable and moving seasonality.
- \( M8 \) to \( M11 \) measure if the movements due to the short-term quasi-random variations and movements due to the long term changes are not changing too much over the years. If the changes are too strong then the seasonal factors could be erroneous. In such a case a correction for a seasonal break or the change of the seasonal filter should be considered.

The \( Q \) statistic is a composite indicator calculated from the \( M \) statistics.

\[
Q = \frac{10M1 + 11M2 + 10M3 + 8M4 + 11M5 + 10M6 + 18M7 + 7M8 + 7M9 + 4M10 + 4M11}{100} \quad [5.22]
\]

\( Q = Q - M2 \) (also called \( Q2 \)) is the \( Q \) statistic for which the \( M2 \) statistics was excluded from the formula, i.e.:

\[
Q - M2 = 10M1 + 10M3 + 8M4 + 11M5 + 10M6 + 18M7 + 7M8 + 7M9 + 4M10 + 4M11 \quad [5.23]
\]

If a time series does not cover at least 6 years, the \( M8, M9, M10 \) and \( M11 \) statistics cannot be calculated. In this case the \( Q \) statistic is computed as:

\[
Q = \frac{14M1 + 15M2 + 10M3 + 8M4 + 11M5 + 10M6 + 32M7}{100} \quad [5.24]
\]

The model has satisfactory quality if the \( Q \) statistic is less than 1.

The tables displayed in the Quality measures → Details node correspond to the F-set of tables produced by the original X-11 algorithm. To facilitate the analysis of the results, the numbers and the names of the tables are given under each table following the convention used in LADIRAY, D., and QUENNEVILLE, B. (1999).

The first table presents the average without regard to sign of the percent changes (multiplicative model) or differences (additive model) over several periods (from 1 to 12 for a monthly series, from 1 to 4 for a quarterly series) for the following series:

\[\text{See 7.6.2.6.}\]
• 0 – Original series (table A1);
• CI – Final seasonally adjusted series (table D11);
• I – Final irregular component (table D13);
• C – Final trend (table D12);
• S – Final seasonal factors (table D10);
• P – Preliminary adjustment coefficients, i.e. regressors estimated by the RegARIMA model (table A2);
• TD&H – Final calendar component (tables A6 and A7);
• Mod. 0 – Original series adjusted for extreme values (table E1);
• Mod. CI – Final seasonally adjusted series corrected for extreme values (table E2);
• Mod. I – Final irregular component adjusted for extreme values (table E3).

In the case of the additive decomposition, for each component the average absolute changes over several periods are calculated as\(^{132}\):

\[
Component_d = \frac{1}{n} \sum_{t=d+1}^{n} |Table_t - Table_{t-d}|
\]

where:
\(d\) – time lag in periods (from a monthly time series \(d\) varies from 1 to 4 or from 1 to 12;
\(n\) – total number of observations per period;
Component – the name of the component;
Table – the name of the table that corresponds to the component.

| Average percent change without regard to sign over the indicated span. |
|-----------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Span | 0 | CI | I | C | S | P | TD&H | Mod. 0 | Mod. CI | Mod. I |
| 1    | 7.50 | 3.61 | 3.49 | 1.42 | 6.99 | 0.00 | 0.00 | 7.75 | 3.57 | 3.29 |
| 2    | 5.33 | 4.60 | 3.90 | 2.00 | 3.57 | 0.00 | 0.00 | 5.40 | 4.61 | 3.55 |
| 3    | 8.23 | 5.75 | 3.74 | 4.39 | 7.16 | 0.00 | 0.00 | 8.53 | 5.50 | 3.39 |
| 4    | 6.36 | 6.75 | 3.76 | 5.84 | 5.00 | 0.00 | 0.00 | 6.74 | 6.74 | 3.56 |

Figure 7.86: Table F2A – changes, in the absolute values, of the principal components.

Next, Table F2B of relative contributions of the different components to the differences (additive model) or percent changes (multiplicative model) in the original series is displayed. They express the relative importance of the changes in each component. Assuming that the components are independent, the following relation is true:

\(^{132}\) For the multiplicative decomposition the following formula is used: Component_d = \(\frac{1}{\sum_{t=d+1}^{n} |Table_t|} \cdot \frac{1}{\sum_{t=d+1}^{n} |Table_t - Table_{t-d}|} \).
\[ O_{d^2} = C_{d^2} + S_{d^2} + I_{d^2} + P_{d^2} + TD\&H_d^2. \] [5.26]

In order to simplify the analysis, the approximation can be replaced by the following equation:

\[ O_{d^2} = C_{d^2} + S_{d^2} + I_{d^2} + P_{d^2} + TD\&H_d^2. \] [5.27]

The notation is the same as for Table F2A. The column Total denotes total changes in the raw time series.

Data presented in Table F2B indicate the relative contribution of each component to the percent \( S_2 \) changes (differences) in the original series over each span, and are calculated as:

\[ \frac{I_{d^2}}{O_{d^2}} \quad \frac{C_{d^2}}{O_{d^2}} \quad \frac{S_{d^2}}{O_{d^2}} \quad \frac{P_{d^2}}{O_{d^2}} \quad TD\&H_d^2 \]

and \( TD\&H_d^2 \) where: \( O_{d^2}^2 = I_{d^2}^2 + C_{d^2}^2 + S_{d^2}^2 + P_{d^2}^2 + TD\&H_d^2 \). The last column presents the Ratio calculated as:

\[ 100 \times \frac{O_{d^2}^2}{O_{d^2}^2}, \] which is an indicator of how well the approximation \( (O_{d^2})^2 \approx O_{d^2}^2 \) holds.

**Relative contributions to the variance of the percent change in the components of the original series.**

<table>
<thead>
<tr>
<th>Span</th>
<th>I</th>
<th>C</th>
<th>S</th>
<th>P</th>
<th>TD&amp;H</th>
<th>Total</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>17.53</td>
<td>3.27</td>
<td>79.20</td>
<td>0.00</td>
<td>0.00</td>
<td>100.00</td>
<td>102.79</td>
</tr>
<tr>
<td>2</td>
<td>37.38</td>
<td>24.71</td>
<td>37.91</td>
<td>0.00</td>
<td>0.00</td>
<td>100.00</td>
<td>115.35</td>
</tr>
<tr>
<td>3</td>
<td>13.97</td>
<td>23.47</td>
<td>62.56</td>
<td>0.00</td>
<td>0.00</td>
<td>100.00</td>
<td>112.79</td>
</tr>
<tr>
<td>4</td>
<td>26.47</td>
<td>73.53</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>100.00</td>
<td>105.49</td>
</tr>
</tbody>
</table>

Figure 7.87: Table F2B – relative contribution of components to changes in the raw series.

For each time lag \( d \), the average and standard deviation of changes are calculated, taking into consideration the sign of the changes and its components.

**Average percent change with regard to sign and standard deviation over indicated span.**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.97</td>
<td>8.67</td>
<td>0.05</td>
<td>3.73</td>
<td>1.41</td>
<td>0.48</td>
<td>0.53</td>
<td>8.01</td>
<td>1.46</td>
<td>3.81</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3.19</td>
<td>5.72</td>
<td>0.15</td>
<td>4.48</td>
<td>2.66</td>
<td>0.97</td>
<td>0.24</td>
<td>4.42</td>
<td>3.02</td>
<td>4.72</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4.97</td>
<td>9.47</td>
<td>0.09</td>
<td>4.52</td>
<td>4.36</td>
<td>1.44</td>
<td>0.55</td>
<td>8.30</td>
<td>4.46</td>
<td>4.57</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5.93</td>
<td>3.81</td>
<td>0.10</td>
<td>4.32</td>
<td>5.50</td>
<td>1.90</td>
<td>0.00</td>
<td>6.01</td>
<td>5.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 7.88: Table F2C – Averages and standard deviations of changes as a function of the time lag.

Average duration of run is an average number of consecutive monthly (or quarterly) changes in the same direction (no change is counted as a change in the same direction as the preceding change). JDemetra+ displays this indicator for the seasonally adjusted series and for the trend and irregular components.
The ratios for each value of time lag \( d \), presented in Table F2E, are computed on a basis of the data in Table F2A.

The relative contribution of components to the variance of the stationary part of the original series is calculated for irregular component (\( I \)), trend made stationary\(^{133} \) (\( C \)), seasonal component (\( S \)) and calendar effects (TD&H). The short description of the calculation method is given in LADIRAY, D., and QUENNEVILLE, B. (1999).

\[
\text{Relative contribution of the components to the stationary portion of the variance in the original series}
\]

The last table shows the autocorrelogram of the irregular component from Table D13. In the case of multiplicative decomposition it is calculated for time lags between 1 and the number of periods per year +2 using the formula\(^{134} \):

\[
Corr_{l} = \frac{\sum_{t=k+1}(I_{t} - 1)(I_{t-k} - 1)}{\sum_{t}(I_{t} - 1)^{2}}
\]

\[^{133}\text{The component is estimated by extracting a linear trend from the trend component presented in Table D12.}\]

\[^{134}\text{For the additive decomposition the formula is } Corr_{l}.\]
\[ r = 1 \]

where \( N \) is number of observations in the time series and \( k \) the lag.

<table>
<thead>
<tr>
<th>Autocorrelation of the irregular</th>
<th>1</th>
<th>-0.601</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>0.200</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.319</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>-0.147</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.187</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>-0.138</td>
</tr>
</tbody>
</table>

Figure 7.92: Table F2G – Autocorrelation of the irregular component.

The Cochran test allows for detection of the heterogeneity of a series of variances. X-13-ARIMASEATS uses this test in the extreme value detection procedure to check if the irregular component is heteroskedastic. The standard errors of the irregular component are then used for an identification of the extreme values. If the null hypothesis that for all the periods (months, quarters) the variances of the irregular component are identical is rejected, the standard errors will be computed separately for each period, in case the option \texttt{Calendarsigma=signif} has been selected.

For each \( i \)th month we will be looking at the mean annual changes for each component by calculating:

\[
S_i = N_{i-1} \sum_{t=2}^{N-1} (S_{i,t} - S_{i,t-1}) \quad \text{and} \quad I_i = \frac{1}{N_i} \sum_{t=2}^{N_i-1} (I_{i,t} - I_{i,t-1})
\]

where \( N_i \) refers to the number of \( i \) months in the data, and the moving seasonality ratio of month \( i \): \( MSR_i = \frac{S_i}{I_i} \). These ratios are published in Table D9A in X13. In JD+ they are presented in the details of the quality measures.

The Moving Seasonality Ratio (MSR) is used to measure the amount of noise in the Seasonal-Irregular component. By studying these values, the user can select for each period the seasonal filter that is most suitable given the noisiness of the series. (see option \texttt{Seasonal filter} in 5.1.2.2). The description of MSR and the rules for selecting the seasonal filters are given in 7.1.3.2.2.
7.2.3. Multi-processing

The Multi processing section, which belongs to the Seasonal adjustment node, is designed to store results of the seasonal adjustment procedure performed with the TRAMO/SEATS or X-13ARIMASEATS methods. The Multi processing is designed for quick and efficient seasonal adjustment of large data sets using different seasonal adjustment methods and different specifications. The functionality is activated from the main menu.

![Figure 7.95: Launching a seasonal adjustment process for a dataset.](image)

The default specification used for multi-documents can be set in the Tools → Options item from the main menu, where the user is expected to specify the seasonal adjustment specification in the Statistics panel.
The consecutive steps necessary to launch the seasonal adjustment process for a dataset are described in the ‘JDemetra+ User Guide’ (2016), case study 3.1.2. The in-depth analysis of seasonal adjustment results and possible ways to correct the deficiencies of the specification and improve the modelling are given in the ‘JDemetra+ User Guide’ (2016), item 3.2.1. Further functionalities intended for the regular production of the seasonally adjusted data, such as summarising the results, saving and refreshing options are discussed in the ‘JDemetra+ User Guide’ (2016), item 3.2.2.

Within one multi-document the user may apply different specifications to the different time series and/or perform the seasonal adjustment for a given time series using different specifications in order to compare the results.

The multi-document window (called the SAProcessing-number window) consists of three panels. The first one – Processing – presents the list of the series. Once the seasonal adjustment process is launched, concise information about the quality of the results is displayed in the Quality column and major problems are manifested in the Warnings column. When the user clicks on an individual time series in the SAProcessing-number window, detailed results are displayed in the panel below the list of series. By default, a summary of results is displayed, accompanied by two graphs: the original data, seasonally-adjusted series and trend on the left and the SI ratios on the right. All the options available for this panel, including launching the seasonal adjustment process, are presented in the ‘JDemetra+ User Guide’ (2016), items 3.1.2 and 3.2.1.
The **Summary** panel gives general information on the results of the adjustment. The report is organised into sections corresponding to the SA methods and frequencies of the time series (monthly and quarterly) which are available from the menu on the top of the panel. The example below shows that the X-13ARIMA-SEATS method has been applied to three monthly time series. All of them have been logarithmically transformed. The list of the ARIMA models shows the model parameters used in the time series set. There were 6 outliers detected, the majority of which were additive outliers. Calendar effects haven’t been detected for any of the time series seasonally adjusted with the X-13ARIMA-SEATS method.
The last panel – *Matrix view* panel – provides information similar to the matrix output of TSW+ (TRAMO/SEATS for the Windows program).

**Figure 7.99: The Matrix panel view.**

The summary information is divided into five folders that are available from the top of the panel:

- **Main** – contains main information about the number of observations in the original time series, the structure of the ARIMA model (number of parameters: $p, d, q, BD, BP, BQ$) used in preprocessing step and its basic properties ($BIC$ – the value of Bayesian information criteria, $SE(\text{res})$ – standard error of residuals, $Q - \text{val}$ - value of the Ljung-Box test on 24 lags).
- **Calendar** – presents calendar specification results, including estimated parameters and their t-value;
- **Outliers** – outlier included in the RegARIMA model including, date, type, parameter value and T-value;
- **Arma** – the values of the autoregressive and moving average parameters and their t-value values;
- **Tests** – results of different tests computed on the residuals including:
  - **Skewness** – the skewness of the distribution of the residuals;
  - **Kurtosis** – the kurtosis of the distribution of the residuals;
  - **Ljung-Box** – the Ljung-Box test on autocorrelation of the residuals up to 24 lags (monthly series or 8 lags (quarterly series));
  - **LB on seas** – the Ljung-Box test on autocorrelation of the seasonal residuals;
  - **LB on sq.** – the Ljung-Box test of linearity of residuals performed on squared residuals;
  - **Bias** - test that compares the annual totals of the original series with those of the seasonally adjusted series;
  - **TD peak** – test for the presence of a significant peak at the calendar frequency in the series spectrum;
- *Seas. Peak* – test for the presence of significant peaks at the seasonal frequencies in the series spectrum;

- *Visual TD peak* – visual test for the presence of a significant peak at the calendar frequency in the series spectrum;

- *Visual seas peak* – visual test for the presence of significant peaks at the seasonal frequencies in the series spectrum.

The matrices can be copied by using the local menu option, and used in other applications, e.g. Excel.
8. Utilities

The Utilities section presents all variables and calendars created by the user. By default, this node contains only a Default calendar (3.3.3) that does not include any public holidays.

![Figure 8.1: The Utilities section.](image1)

8.1. Calendars

The Calendars section stores default and user-defined calendars used for detecting and estimating calendar effects.

![Figure 8.2: The Calendars section with the default calendar and two user-defined calendars.](image2)

Calendar effects are those parts of the movements in the time series that are caused by different number of the weekdays and holidays in calendar months (or quarters). According to the ‘ESS Guidelines on Seasonal Adjustment’ (2015), when time series include calendar effects, the appropriate calendar should be applied to estimate these effects in the seasonal adjustment procedure.

Calendar effects arise as the number of occurrences of each day of the week in a month (or a quarter) differs from year to year. These differences cause regular effects in some series. In particular, such a variation is caused by the leap year effect because of the extra day inserted into February every four years. As with seasonal effects, it is desirable to estimate and remove calendar effects from time series.
Calendar effects can be divided into a mean effect, a seasonal part and a structural part. The mean effect is independent from the period and therefore should be allocated to the trend-cycle. The seasonal part arises from the properties of the calendar that recur each year. For one thing, the number of working days of months with 31 calendar days is on average higher than that of months with 30 calendar days. This effect is part of the seasonal pattern captured by the seasonal component (with the exception of the leap year effect). The structural part of the calendar effect remains to be determined by the calendar adjustment. For example, the number of working days in the same month in different years varies from year to year. This approach is in line with the 'ESS Guidelines on Seasonal Adjustment' (2015).

JDemetra+ estimates calendar effects by adding regressors to the RegARIMA or TRAMO model. The detailed procedure applied in JDemetra+ for the creation of these regression variables is discussed in 7.2. They are used for calculation of the regression variables which are added to the equation estimated by the RegARIMA or the TRAMO procedures.

The list of available calendars is displayed in the Workspace window. Three different types of calendar can be defined:

- National calendars, identified by specific days;
- Composite calendars, defined as a weighted sum of several national calendars;
- Chained calendars, defined by two national calendars and a break date.

Consecutive steps that results in defining each of these calendars, displaying and editing them are presented in the ‘JDemetra+ User Guide’ (2016), item 3.4.3.

8.2. Variables

Variables are simply time series used as explanatory regressors in the RegARIMA and the TRAMO models. Although JDemetra+ allows the user for indicating any time series as a variable to avoid misleading or erroneous results, the following rules should be kept:

- User-defined regression variables are used for measuring abnormalities and therefore they should not contain a seasonal pattern.
- JDemetra+ assumes that user-defined regressors are already in an appropriately centred form. Therefore the mean of each user-defined regressor needs to be subtracted from the regressor or means for each calendar period (month or quarter) need to be subtracted from each of the user-defined regressors.

JDemetra+ considers two kinds of user-defined regression variables:
- Static variables, usually imported directly from external software (by drag/drop or copy/paste);
- Dynamic variables, coming from files that have been opened with the browsers available from the Providers window.

It should be emphasised that JDemetra+ works on the assumption that static variables imported directly from external software (for instance Excel) must be formatted as it has been presented in section 2.1. To import them, go to the Utilities section and then chose Variables in the Workspace window, right-click on it and select New. A new node Vars-1 appears, which can be opened by double clicking. Then drag and drop the time series from an Excel file. The observations for static variables cannot be changed. The only way to update static series is to remove them from the list and to re-import them with the same names.

Dynamic variables are imported into the Variables panel by dragging and dropping series from a browser of the application, available in the Providers window.

By default, JDemetra+ uses the conventions Vars_#number to name the series in the Variables window. The original name of the series is recorded in the Description column of the Variables window. The name of the series in the Variables window can be changed by selecting a series and clicking once again when it has been selected allowing the variable to be renamed.

![Figure 8.3: Local menu for the user-defined variables.](image)

Dynamic variables are automatically updated each time the application is re-opened. Therefore, it is a convenient solution for creating user-defined variables.
9. Annex

9.1. Seasonal adjustment methods – TRAMO/SEATS and X-13ARIMA-SEATS

TRAMO/SEATS is a model-based seasonal adjustment method developed by Víctor Gómez (Ministerio de Hacienda), and Agustín Maravall (Banco de España). It consists of two linked programs: TRAMO and SEATS. TRAMO (“Time Series Regression with ARIMA Noise, Missing Observations, and Outliers”) performs estimation, forecasting, and interpolation of regression models with missing observations and ARIMA errors, in the presence of possibly several types of outlier. SEATS (“Signal Extraction in ARIMA Time Series”) performs an ARIMA-based decomposition of an observed time series into unobserved components. Information about the TRAMO/SEATS method available in this section derives directly from papers by Víctor Gómez and Agustín Maravall; the most important ones are: GÓMEZ, V., and MARAVALL, A. (1996), GÓMEZ, V., and MARAVALL, A. (2001a, b) and MARAVALL, A. (2009). More information about the TRAMO/SEATS method, TRAMO/SEATS software (DOS version and TSW+ – Tramo Seats Windows software and several interfaces) and its documentation as well as papers on methodology and application of the programs, can be found in the dedicated section of the Banco de España site. (www.bde.es, Services → Professionals → Statistical and Econometric Software).

X-13ARIMA-SEATS is a seasonal adjustment program developed and supported by the U.S. Census Bureau. It is based on the U.S. Census Bureau’s earlier X-11 program, the X-11-ARIMA program developed at Statistics Canada, the X-12-ARIMA program developed by the U.S. Census Bureau, and the SEATS program developed at the Banco de España. The program is now used by the U.S. Census Bureau for seasonal adjustment of time series. Users can download the X-13ARIMASEATS application which is a Windows interface for the X-13ARIMA-SEATS program. Detailed information on X-13ARIMA-SEATS can be found at www.census.gov.

In contrast to the earlier product (X-12-ARIMA), X-13ARIMA-SEATS includes not only the enhanced X-11 seasonal adjustment procedure but also the capability to generate ARIMA model-based seasonal adjustment using a version of the SEATS procedure originally developed by Víctor Gómez and Agustín Maravall at the Banco de España. The program also includes a variety of new tools to overcome adjustment problems and thereby enlarge the range of economic time series that can be adequately seasonally adjusted.

In general, X-13ARIMA-SEATS can perform seasonal adjustment in two ways: either using ARIMA model-based seasonal adjustment as in SEATS or by means of an enhanced X-11 method. The ARIMA model-based method implemented in SEATS is characterised in 7.1.2 while the non-parametric method X-11 implemented in X-13ARIMA-SEATS is presented in 7.1.3. The seasonal adjustment methods available in JDemetra+ aim to decompose a time series into components and remove seasonal fluctuations from the observed time series. The X-11 method considers monthly and quarterly series while SEATS is able to decompose series with 2, 3, 4, 6 and
12 observations per year. The main components, each representing the impact of certain types of phenomena on the time series \((X_t)\), are:

- The trend-cycle \((T_t)\) that captures long-term and medium-term behaviour;
- The seasonal component \((S_t)\) representing intra-year fluctuations, monthly or quarterly, that are repeated more or less regularly year after year;
- The irregular component \((I_t)\) combining all the other more or less erratic fluctuations not covered by the previous components.

In general, the trend-cycle consists of 2 sub-components:

- The trend representing the long-term evolution of the series;
- The cycle, that represents the smooth, almost periodic movement around the trend, revealing a succession of phases of growth and recession.

For seasonal adjustment purposes both TRAMO/SEATS and X-13ARIMA-SEATS do not separate the trend from the cycle as these two components are usually too short to perform their reliable estimation. Consequently, hereafter TRAMO/SEATS and X-13ARIMA-SEATS estimate the trend-cycle component. However, the original TRAMO/SEATS may separate trend from cycle through the Hodrick-Prescott filter using the output of the standard decomposition. It should be remembered that JDemetra+ refers to the trend-cycle as trend \((T_t)\), and consequently this convention is used in this document.

TRAMO/SEATS considers two decomposition models:

- The additive model: \(X_t = T_t + S_t + I_t\);
- The log additive model: \(\log(X_t) = \log(T_t) + \log(S_t) + \log(I_t)\).

Apart from these two decomposition types X-13ARIMA-SEATS allows the user to apply also the multiplicative model: \(X_t = T_t \times S_t \times I_t\).

A time series \(x_t\), which is a subject to decomposition, is assumed to be a realisation of a discrete-time stochastic, covariance-stationary linear process, which is a collection of random variables \(\{x_t\}\), where \(t\) denotes time. It can be shown that any stochastic, covariance-stationary process can be presented in the form:

\[
x_t = \mu_t + \tilde{x}_t, \tag{7.1}
\]

where \(\mu_t\) is a linearly deterministic component and \(\tilde{x}_t\) is a linearly interderministic component, such as:
\[ \hat{x}_t = \sum_{j=0}^{\infty} \psi_j a_{t-j}, \]  
[7.2]

where \( \sum_{j=0}^{\infty} \psi_j^2 < \infty \) (coefficients \( \psi_j \) are absolutely summable), \( \psi_0 = 1 \) and \( a_t \) is the white noise error with zero mean and constant variance \( V_a \). The error term \( a_t \) represents the one-period ahead forecast error of \( x_t \), that is:

\[ a_t = \hat{x}_t - \hat{x}_{t-1}. \]  
[7.3]

where \( \hat{x}_{t-1} \) is the forecast of \( \hat{x}_t \) made at period \( t - 1 \). As \( a_t \) represents what is new in \( \hat{x}_t \) in point \( t \), i.e., not contained in the past values of \( \hat{x}_t \), it is also called innovation of the process. From [7.3] \( \hat{x}_t \) can be viewed as a linear filter applied to the innovations.

The equation 7.1 is called a Wold representation. It presents a process as a sum of linearly deterministic component \( \mu_t \) and linearly interderministic component \( \sum_{j=0}^{\infty} \psi_j a_{t-j} \), the first one is perfectly predictable once the history of the process \( x_{t-1} \) is known and the second one is impossible to predict perfectly. This explains why the stochastic process cannot be perfectly predicted.

Under suitable conditions \( \hat{x}_t \) can be presented as a weighted sum of its past values and \( a_t \), i.e.:

\[ \hat{x}_t = \sum_{j=0}^{\infty} \pi_j \hat{x}_{t-j} + a_t. \]  
[7.4]

In general, for the observed time series, the assumptions concerning the nature of the process [7.1] do not hold for various reasons. Firstly, most observed time series display a mean that cannot be assumed to be constant due to the presence of a trend and the seasonal movements. Secondly, the variance of the time series is may not be constant in time. Finally, observed time series usually contain outliers, calendar effects and regression effects, which are treated as deterministic. Therefore, in practice a prior transformation and an adjustment need to be applied to the time series. The constant variance is usually achieved through taking a logarithmic transformation and the correction for deterministic effects, while stationarity of the mean is achieved by applying regular and seasonal differencing. These processes, jointly referred to as preadjustment or linearization, can be performed with the TRAMO or RegARIMA models. Besides the linearisation, forecasts and backcasts of stochastic time series are estimated with the ARIMA model, allowing for later application of linear filters at both ends of time series. The estimation performed with these models delivers the stochastic part of the time series, called the linearised series, which is assumed to be an output of a linear stochastic process.\(^{135}\) The deterministic effects are removed from the time series and used to form the final components.

---

\(^{135}\) MARAVALL, A. (2009).
In the next step the linearised series is decomposed into its components. There is a fundamental difference in how this process is performed in TRAMO/SEATS and X-13ARIMA-SEATS. In TRAMO/SEATS the decomposition is performed by the SEATS procedure, which follows a so called ARIMA model based approach. In principle, it aims to derive the components with statistical models. More information is given in 7.1.2. X-13ARIMA-SEATS offers two algorithms for decomposition: SEATS and X-11. The X-11 algorithm, which is described in 7.1.3, decomposes a series by means of linear filters. Finally, in both methods the final components are derived by the assignment of the deterministic effects to the stochastic components. Consequently, the role of the ARIMA models is different in each method. TRAMO/SEATS applies the ARIMA models both in the preadjustment step and in the decomposition procedure. On the contrary, when the X-11 algorithm is used for decomposition, X-13ARIMA-SEATS uses the ARIMA model only in the preadjustment step. In summary, the decomposition procedure that results in an estimation of the seasonal component requires prior identification of the deterministic effects and their removal from the time series. This is achieved through the linearisation process performed by the TRAMO and the RegARIMA models, shortly discussed in 7.1.1. The linearised series is then decomposed into the stochastic components with SEATS (7.1.2) or X-11 (7.1.3) algorithms.

9.1.1. Linearisation with the TRAMO and RegARIMA models

The primary aim of seasonal adjustment is to remove the unobservable seasonal component from the observed series. The decomposition routines implemented in the seasonal adjustment methods make specific assumptions concerning the input series. One of the crucial ones is that the input series is stochastic, i.e. it is clean of deterministic effects. Another important limitation derives from the symmetric linear filter used in TRAMO/SEATS and X-13ARIMA-SEATS. A symmetric linear filter cannot be applied to the first and last observations with the same set of weights as for the central observations\textsuperscript{136}. Therefore, for the most recent observations these filters provide estimates that are subject to revisions.

To overcome these constrains both seasonal adjustment methods discussed here include a modelling step that aims to analyse the time series development and provide a better input for decomposition purposes. The tool that is frequently used for this purpose is the ARIMA model, as discussed by BOX, G.E.P., and JENKINS, G.M. (1970). However, time series are often affected by outliers, other deterministic effects and missing observations. The presence of these effects is not in line with the ARIMA model assumptions. The presence of outliers and other deterministic effects impede the identification of an optimal ARIMA model due to the important bias in the estimation of parameters of sample autocorrelation functions (both global and partial)\textsuperscript{137}. Therefore, the original series need to be corrected for any deterministic effects and missing observations. This process is called linearisation and results in the stochastic series that can be modelled by ARIMA.

\textsuperscript{136} DAGUM, E.B. (1980).
\textsuperscript{137} GÓMEZ, V., and MARAVALL, A. (2001b). Autocorrelation and partial autocorrelation functions are described in 7.9.
For this purpose both TRAMO and RegARIMA use regression models with ARIMA errors. With these models TRAMO and RegARIMA also produce forecasts.

\[ z_t = y_t \beta + x_t, \quad [7.5] \]

where:
- \( z_t \) is the original series;
- \( \beta = (\beta_1, \ldots, \beta_n) \) – vector of regression coefficients;
- \( y_t = (y_{1t}, \ldots, y_{nt}) \) – \( n \) regression variables (the trading day variables, the leap year effect, outliers, the Easter effect, ramps, intervention variables, user-defined variables);
- \( x_t \) – the disturbance that follows the general ARIMA process: \( \phi(B) \delta(B)x_t = \theta(B)a_t \); \( \phi(B), \theta(B) \) and \( \delta(B) \) are the finite polynomials in \( B \); \( a_t \) is a white-noise variable with zero mean and a constant variance.

The polynomial \( \phi(B) \) is a stationary autoregressive (AR) polynomial in \( B \), which is a product of the stationary regular AR polynomial in \( B \) and the stationary seasonal polynomial in \( B^s \):\(^{138}\)

\[ \phi(B) = \phi_p(B)\Phi_p(B^s) = (1 + \phi_1B + \cdots + \phi_pB^p)(1 + \Phi_1B^s + \cdots + \Phi_pB^{ps}) \quad [7.6] \]

where:
- \( p \) – the number of regular AR terms, (in JDemetra+ \( p \leq 3 \));
- \( p_s \) – the number of seasonal AR terms, (in JDemetra+ \( p_s \leq 1 \));
- \( s \) – the number of observations per year (frequency of the time series).

The polynomial \( \theta(B) \) is an invertible moving average (MA) polynomial in \( B \), which is a product of the invertible regular MA polynomial in \( B \) and the invertible seasonal MA polynomial in \( B \):

\[ \theta(B) = \theta_q(B)\Theta_q(B^s) = (1 + \theta_1B + \cdots + \theta_qB^q)(1 + \Theta_1B^s + \cdots + \Theta_qB^{qs}) \quad [7.7] \]

where:
- \( q \) – the number of regular MA terms, (\( q \leq 3 \));
- \( q_s \) – the number of seasonal MA terms, (\( q_s \leq 1 \)).

The polynomial \( \delta(B) \) is the non-stationary AR polynomial in \( B \) (unit roots):

\(^{138}\) The notation used by TRAMO for the polynomials is different from the one commonly used in the literature, for example in HAMILTON, J.D. (1994) the AR polynomial is denoted as \( \phi(B) = 1 - \phi_1B + \cdots + \phi_pB^p \).
\[ \delta(B) = (1 - B)^d(1 - B^s)^{d_s} \]  

[7.8]

where:

\( d \) – regular differencing order, \( d \leq 1 \);

\( d_s \) – seasonal differencing order, \( d_s \leq 1 \).

JDemetra+ uses notation: \( P, D, Q, BP, BD, BQ \) instead of, respectively: \( p, d, q, p_s, d_s, q_s \). Therefore, the structure of the ARIMA \((p, d, q)(P, D, Q)\) model is denoted in JDemetra+ as ARIMA \((P, D, Q)(BP, BD, BQ)\).

Both TRAMO and X-12-ARIMA allows for automatic identification of an ARIMA model extended for regression variables. The procedure includes test for logarithmic transformation (so-called the test for the log-level specification), selection of the ARIMA model structure and regressors. The estimated deterministic effects are removed from the time series to improve the estimation of the time series components. Forecasts produced by the ARIMA model provide an input for the linear filters used in the decomposition step. In summary, the application of an ARIMA model to the original series vastly improves the quality and stability of the estimated components. The details of the estimation procedure of the ARIMA model with regression variables are presented in the later in this section.

Among the deterministic effects one can distinguish between the calendar effects and outliers. The calendar effects are discussed in 7.2. The impact of different types of outliers on time series is illustrated with several examples by FRANSES, P.H. (1998) For example, it was explained that additive outliers, which are described later in this section, yield large values of skewness and kurtosis, and hence failure of normality test. They also increase the standard error of the estimation parameters. This effect is especially prominent when the size of the outlier is considerable.

KAISER, R., and MARAVALL, A. (1999) express the impact of the outliers on the observed series as:\(^{139}\):

\[ y_t^* = \sum_{j=1}^k \xi_j (B) \omega_j I_t^{(\tau_j)} + y_t \]  

[7.9]

where:

\( y_t^* \) – an observed time series;

\( y_t \) – a series that follows the ARIMA model;

\( \omega_j \) – an initial impact of the outlier at time \( t = \tau_j \);

\( I_t^{(\tau_j)} \) – an indicator variable such that is 1 for \( t = \tau_j \) and 0 otherwise;

\( \xi_j(B) \) – an expression that determines the dynamics of the outlier occurring at time \( t = \tau_j \); 

\( B \) – a backshift operator (i.e. \( B^k X_t = X_{t-k} \)).

The optimal choice of regression variables (and/or intervention variables) requires knowledge from the user about the time series being modelled\(^{140}\). On the contrary, outliers, frequently used in modelling seasonal economic time series, can be automatically detected by JDemetra+ (see 4.1.1.3 for TRAMO and 4.1.2.3 for RegARIMA). The procedure described in GÓMEZ, V., and MARAVALL, A. (2001a) identifies the ARIMA model structure in the presence of deterministic effects. Therefore, the number of identified outliers may depend on the ARIMA model estimated on the series.

The types of outliers that can be automatically identified and estimated by JDemetra+ without any user intervention are:

- **Additive Outlier (AO)** – a point outlier which occurs at a given time \( t_0 \). For the additive outlier \( \xi_j(B) = 1 \), which results in the regression variable:

  \[
  AO_{t_0} = \begin{cases} 
  10 & \text{for } t \neq t_0; \\
  \end{cases}
  \]

- **Level shift (LS)** – a variable for a constant level shift beginning at the given time \( t_0 \). For the level shift \( \xi_j(B) = 1 - B \), which results in the regression variable:

  \[
  LS_{t_0} = \begin{cases} 
  -01 & \text{for } t < t_0; \\
  \end{cases}
  \]

- **Temporary change\(^{141}\) (TC)** – a variable for an effect on the given time \( t_0 \) that decays exponentially over the following periods. For the temporary change \( \xi_j(B) = 1 - \delta B \), which results in the regression variable:

  \[
  TC_{t_0} = \begin{cases} 
  \alpha_0 & \text{for } t < t_0; \\
  \end{cases}
  \]

  where \( \alpha \) is a rate of decay back to the previous level (\( 0 < \alpha < 1 \)).

- **Seasonal outliers (SO)** – a variable that captures an abrupt change in the seasonal pattern on the given date \( t_0 \) and maintains the level of the series with a contrasting change spread over the remaining periods. It is modelled by the regression variable:


\(^{141}\) In the TRAMO/SEATS method this type of outlier is called a transitory change.
\[ S_{01} = \begin{cases} 0 & \text{for } t < t_0 \\ 1 & \text{for } t \geq t_0 \text{ same month/quarter as } t_0 \end{cases} \] \quad [7.13]

where \( s \) is frequency of the time series (\( s = 12 \) for a monthly time series, \( s = 4 \) for a quarterly time series).

The shapes generated by the formulas given above are presented in Figure 9.1.

Figure 9.1: Pre-defined outliers built in to JDemetra+.

Within the RegARIMA model it can be also tested if a series of level shifts cancels out to form a temporary level change effect, which is a permanent level shift spanned between two given dates. It is modelled by the regression variable:

\[ T_{L5}^{(t_0,t_1)} = \begin{cases} 0 & \text{for } t < t_0 \\ 1 & \text{for } t \geq t \geq t_0 \\ 0 & \text{for } t > t_1 \end{cases} \] \quad [7.14]
JDemetra+ also identifies other pre-defined regression variables, for which a necessary input, such as location and process that generates the variable, is provided by the user. This group includes:

- **Ramp** – a variable for a linear increase or decrease in the level of the series over a specified time interval $t_0$ to $t_1$. It is modelled by a regression variable:

  
  $$
  \begin{align*}
  RPt(t_0, t_1) &= \begin{cases} 
  -1 & \text{for } t \leq t_0 \\
  -\frac{0}{t-t} - 1 & \text{for } t_0 < t < t_1 \\
  0 & \text{for } t \geq t_1 
  \end{cases} 
  \end{align*}
  \tag{7.15}
  $$

- Intervention variables which are combinations of five basic structures:\n  - dummy variables;\n  - any possible sequence of ones and zeros;\n  - \[ \frac{1}{(1-\delta \theta)^\gamma} (0 < \delta \leq 1), \] \[ \frac{1}{(1-\delta \phi)^\gamma} (0 < \delta \leq 1), \] \[ \frac{1}{(1-\theta)(1-\phi)} \]

  where $s$ is frequency of the time series ($s = 12$ for a monthly time series, $s = 4$ for a quarterly time series).

The structures considered by intervention variables allow for generation of all pre-defined outliers described by [7.10] - [7.15], as well as more sophisticated effects. An example can be a level shift effect, which is reached after a sequence of damped overshootings and undershootings, presented in Figure 9.2 and denoted there as IV. Another example of an outlier that can be created with intervention variables is a pure seasonal outlier (PSO), which, in contrast to the seasonal outlier described above, does not affect the trend. The set of pure seasonal outliers is used to model the seasonal change of regime (SCR) effect, which describes a sudden and sustained change in the seasonal pattern affecting from $t_0$ (possibly) all seasons of the series. It is defined as:

  
  $$
  \begin{align*}
  SCR_{ij} &= \begin{cases} 
  1 & \text{for } tej \text{ and } t < t_0 \\
  0 & \text{for } t \not\in j \text{ or } t \geq t_0; \\
  -1 & \text{for } tes \text{ and } t < t_0 
  \end{cases} \tag{7.16}
  \end{align*}
  $$
  
  where $j = 1, ..., s - 1$.

---


\(^{143}\) Dummy variable is the variable that takes the values 0 or 1 to indicate the absence or presence of some effect.
9.1.1.1. **Automatic model identification procedure in TRAMO**

An algorithm for Automatic Model Identification in the Presence of Outliers (AMI) implemented in TRAMO is based on TSAY, C. (1986) and CHEN, B.-C., and LIU, L.-M. (1993) with some modifications (see GÓMEZ, V., and MARAVALL, A. (2001)). It iterates between two stages: the first is automatic outlier detection and correction and the second automatic model identification. The parameters of the AMI procedure are described in 4.1.1.5. Unless the parameters are set by the user the program runs with the default values.

The algorithm starts with the identification of the default model and pre-testing procedure, where on the first step a test for a log-level specification is performed. It is based on the maximum likelihood estimation of the parameter \( \lambda \) in the Box-Cox transformation, which is a power transformation such that the transformed values of the time series \( y \) are a monotonic function of the observations, i.e. \( y^\lambda = \{ \lambda = 0 \) \ne 0 \}. First, two Airline models (i.e. ARIMA (0,1,1)(0,1,1)) with a mean \( \log y^\lambda \), \( \lambda = 0 \)

are fitted to the time series: one in logs (\( \lambda = 0 \)), other without logs (\( \lambda = 1 \)). The test compares the sum of squares of the model without logs with the sum of squares multiplied by the square of the geometric mean of the (regularly and seasonally) differenced series in the case of the model in logs.
Logs are taken in the case this last function is the minimum. By default, both TRAMO and X-12ARIMA have a slight bias towards the log transformation.

Next, the test for calendar effects is performed with regressions using the default model for the noise and, if the model is subsequently changed, the test is redone. For seasonal series the default model is the Airline model (ARIMA (0,1,1)(0,1,1)) while for the non-seasonal series ARIMA (0,1,1)(0,0,0) with a constant term is used. The default model, which is used as a benchmark model at some next steps of AMI, is determined by the result of the pre-test for possible presence of seasonality.

Once these pre-tests have been completed, the original series is corrected for all pre-specified outliers and regression effects provided by the user, if any. Next, the order of the differencing polynomial $\delta(B)$ that contains the unit roots is identified and it is decided whether to specify a mean for the series or not. The identification of the ARMA model, i.e. the order of $\phi(B)$ and $\theta(B)$ is performed using the Hannan and Rissanen method by the means of minimising the Bayesian information criterion with some constraints aimed at increasing the parsimony and favouring balanced models. This procedure produces the initial values of the ARIMA model parameters. The search is done sequentially: for the fixed regular polynomials, the seasonal ones are obtained, and vice versa. When the estimated roots of the AR and MA processes are close to each other, the order of the ARIMA model can be reduced. The identification and estimation of the model is carried out using Exact Maximum Likelihood (EML) or the Unconditional Least Squares method.

If the calendar effects were identified in the default model, they are included in a new ARIMA model provided that these effects are significant for this model. The estimated residuals from the modified ARIMA model with fixed parameter estimates and the median absolute deviation of the standard deviation of these residuals are used in the outlier detection procedure. For each observation, $t$-tests are computed for all types of outlier considered in the automatic procedure (AO, TC, LS, SO), following the procedure proposed in CHEN, B.-C., and LIU, L.-M. (1993). The program compares $t$-statistics to a critical value determined by the series length. If there are outliers, for which the absolute $t$-values are greater than a critical value, the one with the greatest absolute $t$ value is selected. After correcting for the identified outlier, the process is started again to test if there is another outlier. The procedure is repeated until for none of the potential outliers, the $t$statistic exceeds the critical value. If outliers are detected, then a multiple regression is

---

145 Parsimonious models are those which have a great deal of explanatory power using a relatively small number of parameters. Balanced models are models for which the order of the combined AR and differencing operators is equal to the order of the combined MA operator (see GÓMEZ, V., and MARAVALL, A. (1997)). A model is said to be more balanced than a competing model if the absolute difference between the total orders of the AR plus differencing and MA operators is smaller for one model than another. For description of the Hannan-Rissanen algorithm see HANNAN, E.J., and RISSANEN, J. (1982), GÓMEZ, V., and MARAVALL, A. (2001b) and 7.1.1.4.
146 See 7.2.1.4.
performed using the Kalman filter and the QR algorithm to avoid (as much as possible) masking effects (i.e. detecting spurious outliers) and to correct for the bias produced in the estimators sequentially obtained. If there are outliers for which the absolute $t$-values are greater than the critical value, the one with the greatest absolute $t$-value is selected and the algorithm continues to the estimation of the ARMA model parameters. Otherwise, the algorithm stops. The estimated residuals from the final ARIMA model are checked for adequacy against the estimated residuals produced by the balanced model. The final model identified by the AMI procedure must show some improvement over the default model in these residual diagnostics; otherwise, the program will accept the default model.

9.1.1.2. Automatic model identification procedure in RegARIMA

The original RegARIMA algorithm developed by the U.S. Census Bureau includes two automatic model selection procedures: `automdl` that is based on TRAMO and `pickmdl` that originates from X11-ARIMA88. The algorithm implementation in JDemetra+ for RegARIMA follows the TRAMO logic. It is very similar to the TRAMO procedure presented in the previous section, but contains modifications to make use of the X-13ARIMA-SEATS estimation procedure, which is different from the one that TRAMO uses. The examples of extensions that are specific to RegARIMA only are: special treatment of the leap year effect in the multiplicative model, automatic detection of the length of the Easter effect, option to reduce a series of level shifts to the temporary level shift. In comparison with TRAMO, there are also differences in the values of the default parameters. Besides, by default, RegARIMA does not favour balanced models. The log/level test in RegARIMA is based on the Corrected Akaike Information Criterion (AICC). Similarly, the estimation of calendar effects is based on AICC. As a result, the model selected by RegARIMA can differ from the model that TRAMO would select, especially in the case of series contaminated with deterministic effects and/or those, which are modelled with the mixed ARIMA models.

9.1.1.3. Model selection criteria

Model selection criteria are statistical tools for selecting the optimal order of the ARIMA model. The basic idea behind all these criteria is to obtain as much explanatory power (measured by the value of the likelihood function) with only a few parameters. The model selection criteria essentially choose the model with the best fit, as measured by the likelihood function, and it is a

---

147 See 7.1.1.6.
153 The pre-tested options are: one, eight, and fifteen days before Easter.
154 See 7.1.1.2.
subject to a penalty term, to prevent over-fitting that increases with the number of parameters in the model. Some of the most well-known information criteria are: Akaike Information Criterion (AIC), Corrected Akaike Information Criterion (AICC), HQ Information Criterion (HannanQuinn) and Schwarz-Bayes Information Criterion (BIC).

The formulae for the model selection criteria used by JDemetra+ are:

\[
AIC_N = -2L_N + 2n_p
\]

\[
AICC_N = -2L_N + 2n_p(1 - n_p/2N+N+1)-1,
\]

\[
HQ_N = -2L_N + 2n_p\log\log N,
\]

\[
BIC_N = -2L_N + n_p\log N.
\]

where:

- \(N\) – number of observations in time series;
- \(n_p\) – number of estimate parameters; \(L_N\) – loglikelihood function.

For each model selection criteria the model with the smaller value is preferred. As it can be shown that AIC is biased for small samples, it is often replaced by AICC. To choose the ARIMA model parameters the original RegARIMA uses AICC while TRAMO uses BIC with some constraints aimed at increasing the parsimony and favouring balanced models. The automatic model identification methods implemented in JDemetra+ mostly use BIC, however other criteria are used as well. It should be noted that the BIC criterion imposes a greater penalty term than AIC and HQ, so that BIC tends to select simpler models than those chosen by AIC. The difference between both criteria can be huge if \(N\) is large.

---

9.1.1.4. Hannan-Rissanen algorithm

The Hannan-Rissanen algorithm\textsuperscript{160} is a penalty function method based on the BIC criterion, where the estimates of the ARMA model parameters are computed by means of a linear regression. The Hannan-Rissanen procedure operates on a stationary transformation of the original series, therefore only AR and MA parameters are being identified.

In the first step, a high-order \textit{AR}(m), where \( m > \max(p, q) \), model is fitted to the time series \( X_t \). The maximum order of \( p \) and \( q \) is given a-priori. In JDemetra+ it is equal to 3 for both \( p \) and \( q \). The residuals \( \hat{a}_k \) from this model are used to provide estimates of the innovations in the ARMA model \( a_t \):

\[
a_t = X_t - \sum_{k=1}^{m} \hat{a}_k X_{t-k}.
\]

In the second step the parameters \( p \) and \( q \) of the ARMA model are estimated using a least squares linear regression of \( X_t \) onto \( X_{t-1}, \ldots, X_{t-p} a_{t-1}, \ldots, a_{t-q} \) for a combination of values \( p \) and \( q \). The first step of the procedure is skipped if the ARMA model fitted in the second step includes only an autoregressive part.

Finally, the Hannan-Rissanen algorithm selects a pair of \( p \) and \( q \) values for which \( BIC_{p,q} \) is the smallest. \( BIC_{p,q} \) is defined as:

\[
BIC_{p,q} = \log(\hat{\sigma}_{p,q}^2) + \left( \frac{p+q}{n} \right) \log(n-d),
\]

where:
- \( \hat{\sigma}_{p,q}^2 \) is the maximum likelihood estimator of \( \sigma^2 \);
- \( n - d \) is the number of series in the (regularly and/or) seasonally differenced series.

The advantage of the Hannan-Rissanen algorithm is the speed of computation in comparison with exact likelihood estimation.

9.1.1.5. Initial values for ARIMA model estimation

By default, the initial parameter value in X-13ARIMA-SEATS is 0.1 for all AR and MA parameters. For the majority of time series this default value seems to be appropriate. Introducing better initial values (as might be obtained, e.g., by first fitting the model using conditional likelihood) could slightly speed up convergence. Users are allowed to introduce manually initial values for AR and MA parameters that are then used to start the iterative likelihood maximization. This is rarely

necessary, and in general not recommended. A possible exception to this occurs if the initial estimates that are likely to be extremely accurate are already available, such as when one is re-estimating a model with a small amount of new data added to a time series. However, the main reason for specifying initial parameter values is to deal with convergence problems that may arise in difficult estimation situations\textsuperscript{161}.

### 9.1.1.6. Cancellation of AR and MA factors

A cancellation issue consists in cancelling some factors on both sides of the ARIMA model. This problem concerns the mixed ARIMA \((p, d, q)(P, D, Q)\) models (i.e. \(p > 0\) and \(q > 0\), or \(P > 0\) and \(Q > 0\)). For example, a cancellation problem occurs with ARMA \((1,1)\) model, \((1 - \phi B)z_t = (1 - \theta B)\alpha_t\) when \(\phi = \theta\) as then model is simply of the form: \(z_t = \alpha_t\). Such model causes problems with convergence of the nonlinear estimation. For this reason the X-13ARIMA-SEATS and TRAMO/SEATS programs deal with cancellation problem by computing zeros of the AR and MA polynomials. As the cancellation does not need to be exact, the cancellation limit can be provided by the user\textsuperscript{162}.

\begin{footnotesize}
\textsuperscript{162} Ibid.
\end{footnotesize}
9.1.1.7. Least squares estimation by means of the QR decomposition

We consider the regression model:

\[ y = X\beta + \epsilon. \]  \hspace{1cm} [7.23]

The least squares problem consists in minimizing the quantity \( \|X\beta - y\|^2 \).

Provided that the regression variables are independent, it is possible to find an orthogonal matrix \( Q \) so that \( Q^t X = \begin{pmatrix} R \\ 0 \end{pmatrix} \) where \( R \) is upper triangular.

We have now to minimize:

\[ \|QX\beta - Qy\|^2 = \|Q(0)\beta - Qy\|^2 = \|R\beta - a\|^2 + \|b\|^2\]  \hspace{1cm} [7.24]

where \((Qy)_{0...x-1} = a\) and \((Qy)_{x...n-1} = b\).

The minimum of the previous norm is obtained by setting \( \beta = R^{-1}a \). In that case \( \|R\beta - a\|^2 = 0 \).

The residuals obtained by this procedure are then \( b \), as defined above.

It should be noted that the QR factorization is not unique, and that the final residuals also depend on the order of the regression variables (the columns of \( X \)).

9.1.2. SEATS

SEATS is a program for the estimation of stochastic unobserved components in stochastic time series. It follows a so-called ARIMA-model-based (AMB) methodology for time series decomposition, developed from the work of CLEVELAND, W.P., and TIAO, G.C. (1976), BURMAN, J.P. (1980), HILLMER, S.C., and TIAO, G.C. (1982), BELL, W.R., and HILLMER, S.C. (1984), and MARAVALL, A., and PIERCE, D.A. (1987). Applied in SEATS this methodology aims for identification of the model that provides a satisfactory decomposition e.g. the one that results in a more stable seasonal component, i.e. the one which is estimated with more precision and is subject to smaller revisions.

The input for the model based signal extraction procedure is provided by TRAMO and includes the ARIMA model for the stochastic (linearised) time series \( x_t \) and identified deterministic effects. The procedure consists in an estimation of time series components by means of the Wiener-
Kolmogorov filters as the Minimum Mean Square Error estimators using the UCARIMA (Unobserved Component ARIMA) models.

One of the fundamental assumptions made by SEATS is that the stochastic time series \( x_t \) follows the ARIMA model:

\[
\phi(B) \delta(B) x_t = \theta(B) a_t, \tag{7.25}
\]

where:
- \( \delta(B) \) – a non-stationary autoregressive (AR) polynomial in \( B \) (unit roots);
- \( \theta(B) \) – an invertible moving average (MA) polynomial in \( B \), which is a product of an invertible regular MA polynomial in \( B \) and an invertible seasonal MA polynomial in \( B^S \);
- \( \phi(B) \) – a stationary autoregressive (AR) polynomial in \( B \) and stationary seasonal polynomial in \( B^S \);
- \( a_t \) – a white-noise variable with the variance \( V(a) \).

It is convenient to present (7.25) as:

\[
 z(B) x_t = \theta(B) a_t, \tag{7.26}
\]

where:
- \( z(B) = \phi(B) \delta(B) \) is a product of the stationary and the nonstationary AR polynomials.

The stochastic time series can be predicted using its past observations and making an error. An unpredictable at \( t - 1 \) part of \( x_t \) is called an innovation\(^{163}\) and it is expressed by residuals \( a_t \), which can be also regarded as estimators of the one-period-ahead forecast error of the observed series \( x_t \).

### 9.1.2.1. Derivation of the models for the components

SEATS decomposes the series \( x_t \) received from TRAMO into \( k \) components:

\[
x_t = \sum_{i=1}^{k} x_{it}, \tag{7.27}
\]

\(^{163}\) The definition of innovations is given earlier in the 7.1.
where \( i \) refers to the trend, seasonal, transitory or irregular component. Apart from the irregular component, assumed to be white noise, it is assumed that each component is the outcome of the linear stochastic process, which, using the same notation of [7.25], can be represented as:

\[
z_i(B) x_{it} = \theta_i(B) a_{it}
\]  

[7.28]

The white-noise variable \( a_{it} \) has a zero mean and a constant variance \( V(a_i) \); for \( i \neq j \) \( a_{it} \) is not correlated with \( a_{jt} \) for any \( t \). In an unobserved-components model, residuals \( a_{it} \) are disturbances associated with the unobserved components. These disturbances are functions of the innovations in the series and are called "pseudo-innovations" as they refer to the components that are never observed\(^{164}\). JDemetra+ uses term "innovations" to refer to "pseudo-innovations". It is also assumed that \( z_i(B) \) polynomials are prime and the \( \theta_i(B) \) polynomials do not share unit roots in common.

For each \( i \) the polynomials \( \phi_i(B) \), \( \delta_i(B) \) and \( \theta_i(B) \) are prime and of finite order. The roots of \( \delta_i(B) \) lies on the unit circle; those of \( \phi_i(B) \) lie outside while all the roots of \( \theta_i(B) \) are on or outside the unit circle. Since different roots of the polynomial induce peaks in the spectrum\(^{165}\) of the series for different frequencies, and given that different components are associated with spectral peaks for different frequencies, it is assumed that for \( i \neq j \) the polynomials \( \phi_i(B) \) and \( \phi_j(B) \) do not share any common roots. Finally, it is assumed that the polynomials \( \theta_i(B), i = 1, \ldots, k \) share no unit roots in common. This requirement guarantees invertibility of the overall series because non invertibility is associated with a spectral zero. Therefore, when the polynomials \( \theta_i(B), i = 1, \ldots, k \) share no unit roots in common, there is no frequency for which all component spectra become zero.\(^{176}\)

The equations [7.27] and [7.28] together with associated assumptions define an unobserved component ARIMA (UCARIMA). It should be stressed that, in general, for a given ARIMA model there is no unique UCARIMA representation that generates it.\(^{166}\) For decomposition purpose SEATS makes several assumptions which will be brought out in the subsequent parts of the description of the procedure.

From [7.26] - [7.28] it can be shown that the AR polynomial in the model for \( x_t \) satisfies:

\[
z(B) = \prod_{i=1}^{k} z_i(B)
\]  

[7.29]

and the MA polynomial can be obtained from the relationship

\[
k
\]


\(^{165}\) For description of the spectrum see 7.3.\(^{176}\)


\(^{166}\) Ibid.
\[ \theta(B) a_t = \sum_{i=1}^{k} z_{ni}(B) \theta_i(B) a_{it} \]  

where \( z_{ni}(B) \) is the product of all \( z_j(B), j = 1, ..., k, \) not including \( z_i(B) \).

The decomposition is obtained from the identity

\[
\frac{\theta(B)}{z(B) a_t} = \sum_{i=1}^{k} \frac{\theta_i(B)}{a_{it}} + \frac{u_t}{z(B)} \]  

where \( a_t \) denotes the trend-cycle, seasonal and transitory components and \( u_t \) denotes the irregular component. The AR polynomials for the components \( z_i(B) \) are obtained by factorization of the AR polynomial of the model for \( x_t \). Such decomposition requires calculation of the inverse roots of \( z(B) = 0 \). To understand how SEATS factorizes the AR polynomials, first a concept of root will be explored\(^{167}\).

The solutions of the equation \( 1 + \phi_1 B + \cdots + \phi_p B^p \) can be found by replacing \( B \) by \( z^{-1} \); \( z \neq 0 \), (and multiplying by \( z^p \)) resulting in a so-called characteristic equation:

\[
z^p + \phi_1 z^{p-1} + \cdots + \phi_{p-1} z + \phi_p = 0. \]  

Equation [7.32] has real and complex roots (solutions). A complex number, \( x = a + bi \), with \( a \) and \( b \) both real numbers, can be represented as \( x = r ( \cos \omega + i \sin \omega ) \), where \( i \) is the imaginary unit \( (i^2 = -1) \), \( r \) is the modulus of \( x \), that is \( r = |x| = \sqrt{a^2 + b^2} \), and \( \omega \) is the argument (frequency). When roots are complex, they are always in pairs of complex conjugates. Both the representations of the complex number \( x = a + bi \) have a geometric interpretation with the complex plane established by the real axis and the orthogonal imaginary axis.

Figure 9.3: Geometric representation of a complex number and of its conjugate.

Representing the roots of the characteristic equation in the complex plane allows understanding of how they are allocated to the components. When their modulus $r$ is greater than 1, the solution of the characteristic equation has a systematic explosive behaviour, which is contrary to the developments that can be identified in actual economic series. Therefore, the models estimated by TRAMO/SEATS (and X-13ARIMA-SEATS) never have roots for which the modulus is greater than 1.

The characteristic equations associated to the regular and the seasonal differences, have roots with modulus $r = 1$. They are called non-stationary roots and can be represented on the unit circle, as it shown in Bląd! Nie można odnaleźć źródeł odwołania.. For the seasonal differencing operator applied to a quarterly time series $(1 - B^4)$ the characteristic equation is $z^4 - 1 = 0$ with solutions given by $z = \sqrt[4]{1}$, which imply that $z_1 = 1$, $z_2 = -1$, $z_3 = i$ and $z_4 = -i$. The first two solutions are real and the last two are complex conjugates. They are represented on the unit circle of Figure 7.4.
In the case of the seasonal differencing operator \((1 - B^{12})\) applied to the monthly time series the characteristic equation \(z^{12} - 1 = 0\) has twelve non-stationary solutions given by \(z = \frac{12}{12};\) two real and ten complex conjugates.

The complex conjugates roots generate the periodic movements of the type:

\[
z_t = A \cos(\omega t + B).
\]

where:
- \(A\) – amplitude;
- \(\omega\) – angular frequency (in radians); \(B\)
- \(B\) – phase (angle at \(t = 0\)).

The frequency \(f\), i.e. the number of cycles per unit time, is \(\frac{\omega}{2\pi}\). If it is multiplied by \(s\), the number of observations per year, the number of cycles completed in one year is derived. Therefore, for monthly time series the seasonal movements are produced by complex conjugates roots at the angular frequencies: \(\frac{\pi}{12}, \frac{2\pi}{12}, \frac{5\pi}{12}\), and \(\pi\). The corresponding number of cycles per year and the length of the movements are presented in Bląd! Nie można odnaleźć źródła odwołania.
Table 9.1: Seasonal frequencies for a monthly time series.

<table>
<thead>
<tr>
<th>Angular frequency $\omega$</th>
<th>Frequency (cycles per unit time) $f$</th>
<th>Cycles per year</th>
<th>Length of the movement measured in months</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\pi/6$</td>
<td>0.083</td>
<td>$\pi/6$</td>
<td>12</td>
</tr>
<tr>
<td>$\pi/3$</td>
<td>0.167</td>
<td>$\pi/3$</td>
<td>6</td>
</tr>
<tr>
<td>$\pi/2$</td>
<td>0.250</td>
<td>$\pi/2$</td>
<td>4</td>
</tr>
<tr>
<td>$2\pi/3$</td>
<td>0.333</td>
<td>$2\pi/3$</td>
<td>3</td>
</tr>
<tr>
<td>$5\pi/6$</td>
<td>0.417</td>
<td>$5\pi/6$</td>
<td>2.4</td>
</tr>
<tr>
<td>$\pi$</td>
<td>0.500</td>
<td>$\pi$</td>
<td>2</td>
</tr>
</tbody>
</table>

In case of quarterly series the seasonal movements are produced by complex conjugates roots with $\pi$ angular frequencies at (one cycle per year) and $\pi$ (two cycles per year). 2

SEATS assigns the roots of AR full polynomial to the components according to their associated modulus and frequency, i.e.: \(^{168}\)

- Roots of $(1 - B)^d$ are assigned to trend component.
- Roots of $(1 - B^s)^d_s = ((1 - B)(1 + B + \ldots + B^{s-1}))^{d_s}$ are assigned to the trend component (root of $(1 - B)^{d_s}$) and to the seasonal component (roots of $(1 + B + \ldots + B^{s-1})^{d_s}$).
- When the modulus of the inverse of a real positive root is greater than $k$, then the root is assigned to the trend component, where $k$ is the threshold value controlled by the $rmod$ parameter (in the JDemetra+ specifications this argument is called the Trend boundary)\(^{169}\). Otherwise it is integrated in the transitory component.
- Real negative inverse roots of $\phi(B)$ associated with the seasonal two-period cycle and complex roots for which the argument (angular frequency) is close enough to the seasonal frequency are assigned to the seasonal component if their moduli are greater than $k$. Otherwise they are assigned to the transitory component. Closeness is controlled by the $epsphi$ parameter \(^{170}\) (in the JDemetra+ specifications this argument is called Seasonal tolerance).
- If $d_s$ (seasonal differencing order) is present and $B_{phi} < 0$ ($B_{phi}$ is the estimate of the seasonal autoregressive parameter), the real positive inverse root is assigned to the trend component and the other $(s - 1)$ inverse roots are assigned to the seasonal component. When $d_s = 0$, the root is assigned to the seasonal when $B_{phi} < -0.2$ and/or the overall test

\(^{168}\) For details see MARAVALL, A., CAPORELLO, G., PÉREZ, D., and LÓPEZ, R. (2014).

\(^{169}\) See 5.1.1.2. In JDemetra+ this argument is called Trend boundary.

\(^{170}\) See 5.1.1.2.
for seasonality indicates presence of seasonality. Otherwise it goes to the transitory component. Also, when $Bp_{hi} > 0$, roots are assigned to the transitory component.

It should be highlighted that when $Q > P$, the SEATS decomposition yields a pure MA $(Q - P)$ component (hence transitory). In this case the transitory component will appear even when there is no AR factor allocated to it.

Once these rules are applied, the factorisation of the AR polynomial presented by [7.29] yields to the identification of the polynomials for the components which contain, respectively, the AR roots associated with the trend component, the seasonal component and the transitory component. The AR roots close to or at the trading day frequency generates a stochastic trading day component.\textsuperscript{171}

The MA polynomials and the innovation variances cannot be identified from the UCARIMA model. This underidentification problem can be solved by assuming that for each component the order of MA process is no greater than the order of AR process. In addition to this, to achieve a unique UCARIMA representation of ARIMA model for $x_t$, all additive white noise is added to the irregular component. SEATS requires that the MA polynomial is invertible, which means that it can be expressed as a linear combination of time series lagged values plus the contemporaneous error term. In general, MA($q$) models are called invertible when the solutions to characteristic equation:

\[ 1 + \theta_1 z + \ldots + \theta_q z^q = 0 \quad [7.34] \]

are all outside the unit circle.\textsuperscript{172}

Once the roots are assigned to the components, a partial fraction expansion yields the following decomposition of $x_t$:

\[
\begin{align*}
\theta(B) & = \theta_{\text{trend}}(B) + \theta_{\text{seasonal}}(B) + \theta_{\text{transitory}}(B) \\
\delta(B)\phi(B)u_t = \delta_{\text{trend}}(B)\phi_{\text{trend}}(B)u_t + \delta_{\text{seasonal}}(B)\phi_{\text{seasonal}}(B)u_t + \delta_{\text{transitory}}(B)\phi_{\text{transitory}}(B)u_t + u_t
\end{align*}
\quad [7.35]
\]

where:

$u_t$ – white noise.

\textsuperscript{171} A stochastic trading day component is always modelled as a stationary ARMA(2,2), where the AR contains the roots close to the TD frequency, and the MA(2) is obtained from the model decomposition (MARAVAL, A., and PÉREZ, D. (2011)). This component, estimated by SEATS, is not implemented by the current version of JDemetra+.

\textsuperscript{172} FRANSES, P.H. (1998).
\( a_{\text{trend}, t}, a_{\text{seasonal}, t}, a_{\text{transitory}, t} \) – the uncorrelated innovation in, respectively, the trend, seasonal and transitory components.

\[
\begin{align*}
\theta_{\text{trend}}(B) a_{\text{trend}, t} & \quad \text{model for the trend;} \\
\delta_{\text{trend}}(B) \phi_{\text{trend}}(B) & \quad \theta_{\text{seasonal}}(B) \arrowvert \delta_{\text{seasonal}}(B) \phi_{\text{seasonal}}(B) a_{\text{seasonal}, t} \quad \text{model for the seasonal component;} \\
\delta_{\text{transitory}}(B) \phi_{\text{transitory}}(B) & \quad \theta_{\text{transitory}}(B) a_{\text{transitory}, t} \quad \text{model for the transitory component.}
\end{align*}
\]

The resulting decomposition is called canonical. For this decomposition all components except from the irregular have a spectral minimum of zero and are thus invertible.\(^{173}\) It has been shown by that the canonical decomposition produces as stable as possible (given the stochastic features of the series) trend and seasonal components as it maximizes the variance of the irregular and minimizes the variance of other components\(^{174}\). It can be shown that any other admissible component (i.e. the component for which spectrum is non-negative for all frequencies) is equal to the canonical one plus added noise, and hence the canonical requirement makes the component as smooth as possible\(^{175}\). The example of the pseudo-spectrum for the linearised series is presented in

\[\text{Bląd! Nie można odnaleźć źródła odwolania.}\]


\(^{175}\) Ibid.
Figure 9.5: A comparison of the canonical trend and admissible trend.

However, there is a price to be paid as canonical components can produce large revisions in the preliminary estimators of the components\(^{176}\).

The partial fraction decomposition is performed in the frequency domain. In essence, it consists in portioning of the pseudo-spectrum\(^{177}\) of \(x_t\) into additive spectra of the components. The pseudo-spectrum of \(x_t\) is defined as the Fourier transform of \(\psi(\theta(B)) \psi(\phi(F)) V(a_i)\), where \(\psi_i(F) = \phi^{\theta_i(F)} \psi_i(B)\).

\[ \psi_i(B) = \theta_i(B), \quad B \text{ is a backward operator and } F \text{ is the forward operator.} \]

The pseudo-spectrum is denoted by \(g_i(\omega)\), where \(\omega\) is a frequency.

The example of the pseudo-spectrum for a linearised series is presented in Błąd! Nie można odnaleźć źródła odwołania.. The frequency \(\omega = 0\) is associated with a trend, while for the frequencies

\[ \pi < \frac{\pi}{a} \text{ in the range } [0 + \epsilon_1, -\epsilon_2] \text{ with } \epsilon_1, \epsilon_2 > 0 \text{ and } \epsilon_1 < -\epsilon_2 < \frac{\pi}{2} \text{ the associated period will be longer than } \pi \text{ a year and bounded}^{178}. \]

The frequencies in the range \([, \pi]\) are associated with the short term movements, with the cycle completed in less than a year. If a series contains an important component for a certain frequency, its spectrum should reveal a peak around that frequency caused by the unit AR root. Therefore for a trend the spectral peak should occur at the frequency \(\omega = 0^{179}\) and significant seasonality should result in peaks at seasonal frequencies (\(\frac{\pi}{6}, \frac{\pi}{6}, \frac{\pi}{6}, \frac{\pi}{6}, \pi\) for a monthly time series and \(\frac{\pi}{2}, \pi\) for a quarterly series).

\(^{176}\) MARAVALL, A. (1986).

\(^{177}\) Term pseudo-spectrum is used for non-stationary time series, while term spectrum – for stationary time series. For description of spectrum see 7.3.

\(^{178}\) Ibid.

In the decomposition procedure the pseudo-spectrum of a linearised time series is divided into spectra of four components (Błąd! Nie można odnaleźć źródła odwołania.).

If spectra for all components are non-negative for all frequencies, the decomposition is called an admissible decomposition. The decomposition performed on this step is vital for the estimation of the components because the properties of the component estimator will depend on the admissible
decomposition selected. If the ARIMA model selected by TRAMO leads to non-admissible decomposition, SEATS changes the ARIMA model and uses the modified one to decompose the series. There are also other, rare situations when the ARIMA model chosen by TRAMO is changed by SEATS. It happens when, for example, the ARIMA model generates unstable seasonality or produces a senseless decomposition. Such examples are discussed by MARAVALL, A. (2009).

In the original TRAMO/SEATS the coherence between forecasts of TRAMO and SEATS is guaranteed. In JDemetra+ however, in cases when the ARIMA model is changed by SEATS, the forecasts of raw and linearised series are calculated using the ARIMA model selected by TRAMO, while the decomposition and the components’ forecasts are based on ARIMA model changed by SEATS, constrained to aggregate into the TRAMO forecasts.

9.1.2.2. Estimation of the components with the Wiener-Kolmogorow filter

As the time series components are never observed, their estimators have to be used. For a particular realisation of \(X_T = [x_1, x_2, \ldots x_T]\), where \(T\) is the last observed period, SEATS aims to obtain for each component (except from the irregular) the estimator \(\hat{x}_{it}\) such that \(E[(x_{it} - \hat{x}_{it})^2|X_T] \) is minimized.

This assumption yields the Minimum Mean Square Error (MMSE) estimator. Under the joint normality assumption \(\hat{x}_{it}\) is also equal to the conditional expectation \(E(x_{it}|X_T)\), so it can be presented as a linear function of the elements in \(X_T\):

\[
\hat{x}_{it} = \cdots + v \cdot x_{t-k}|T + \cdots + v \cdot x_{t}|T + \cdots + v \cdot x_{t+k}|T + \cdots \tag{7.36}
\]

In case of an infinite realisation of the series \(x_t\), denoted as \(X\), the optimal estimator of the component \(i\) is given by:

\[
\hat{x}_{it} = \hat{x}_{it}|\infty = E(x_{it}|X) \tag{7.37}
\]

The conditional expectation can be computed by the Wiener-Kolmogorow (WK) filter given by

\[
\hat{x}_{it} = k \cdot \psi \cdot (B \psi^{-1})(F \psi^{-1}) \cdot x_t \tag{7.38}
\]

where \(\psi(B) = \phi(B)\), \(F = B^{-1}\) and \(k = v^\psi(a^\psi)\), \(V(a)\) - a variance of \(a\) and \(V(a)\) - a variance of \(a_t\).

---

\[\text{GÓMEZ, V., and MARAVALL, A. (2001a).}\]
Once the $\psi(B)$ and $\psi(F)$ polynomials are replaced by their rational expressions and common factors are cancelled, [7.37] becomes:

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

[7.39]

where $v_i(B,F) = v_0 + \sum_{j=1}^{\infty} v_{ij} (B^j + F^j)$ is the WK filter.

Unlike the ad-hoc filters, the WK filter applied to the $i$–$th$ component derives from the observed model, which in turn depends on the observed series and consequently the WK filters adapt to the observed series. As a result, the properties of the component estimator will depend on the admissible decomposition selected\textsuperscript{181}. The WK filter is symmetric and centred in both $B$ and $F$ which allows the phase effect to be avoided. When the observed model is not a pure AR model then the WK filter is of infinite length.

By construction, the WK filter adapts itself to the series under consideration, and this adaptability avoids the dangers of under and overestimation with ad-hoc filtering. For example, for the series with a highly stochastic seasonal the filter adapts to the width of the seasonal peaks and the seasonally adjusted series does not display any spurious seasonality\textsuperscript{182}.

The example of the WK filters obtained for a particular time series is shown in Figure 9.8:

WK filters for components.

---


\textsuperscript{182} MARAVALL, A. (1995).
Final estimators for the components

In theory, the derivation of the components requires an infinite realization of $x_t$ in the direction of the past and of the future. However, it can be shown that the WK is convergent, which guarantees that, in practice, it could be approximated by a truncated (finite) filter and, in most applications, for large $k$ the estimator for the central periods of the series can be safely seen as generated by the WK filter:\(^{183}\)

\[
\hat{x}_{it} = \nu_k x_{t-k} + \cdots + \nu_1 x_t + \cdots + \nu_k x_{t+k}
\]  

[7.40]

When $T > 2L + 1$, where $T$ is the last observed period, and $L$ typically expands between 3 and 5 years, the estimator expressed by [7.38] can be assumed as the final (historical) estimator for the central observations of the series\(^{184}\). In practice, the Wiener-Kolmogorow filter is applied to $x_t$ extended with forecasts and backcasts from the ARIMA model.

The estimation procedure of the components will be explained for the two component model (the seasonally adjusted series ($s_t$), called signal and the seasonal component, i.e. noise ($n_t$)). In the


frequency domain Wiener-Kolmogorow filter \( \nu(B, F) \) that provides the final estimator of \( s(t) \) is expressed as the ratio of the \( s(t) \) and \( x(t) \) pseudo-spectra:

\[
\hat{\nu}(\omega) = \frac{g_s(\omega)}{g_x(\omega)}
\]  

[7.41]

The function \( \hat{\nu}(\omega) \) is also referred as the gain of the filter.\(^{185}\) GÓMEZ, V., and MARAVALL, A. (2001a) show that when for some frequency signal (the seasonally adjusted series) dominates the noise (seasonal fluctuations) \( \hat{\nu}(\omega) \) approaches 1. On the contrary, when for some frequency noise dominates the signal \( \hat{\nu}(\omega) \) approaches 0.

The spectrum of the estimator of the seasonal component is expressed as:

\[
g_s(\omega) = \left( \frac{g_s(\omega)}{g_x(\omega)} \right)^2
\]  

[7.42]

where \( \hat{\nu}(\omega) = \frac{[g_s(\omega)]}{g_s(\omega)} \) and \( r(\omega) = \frac{g_x(\omega)}{g_s(\omega)} \) represents the signal-to-noise ratio.

For each \( \omega \) MMSE estimation computes the signal-to-noise ratio. It this ratio is high, then the contribution of that frequency to the estimation of the signal will be also high. Assume that the trend is a signal that need to be extracted from a seasonally time series. Then \( R(0) = 1 \) and the frequency \( \omega = 0 \) will only be used for trend estimations. For seasonal frequencies \( R(\omega) = 0 \), so that these frequencies are ignored in computing the trend resulting in spectral zeros in \( g_s(\omega) \). For this reason, unlike the spectrum of the component, the component spectrum contains dips as it can be seen in the Figure 9.9: Component spectrum and estimator spectrum for trend.

---

\(^{185}\) GÓMEZ, V., and MARAVALL, A. (2001a).
From the equation [7.42] it is clear that the squared gain of the filter determines how the variance of the series contributes to the variance of the seasonal component for the different frequencies. Since \( r(\omega) \geq 0 \), then \( \tilde{v}(\omega) \leq 1 \) and from [7.42] it can be derived that \( g_s(\omega) = \tilde{v}(\omega)g_s(\omega) \), resulting in the observation that the estimator will always underestimate the component. It can be seen that \\

\[ g_s(\omega) = \frac{V(s)}{g_s(\omega)} = \frac{V(s)}{g_s(\omega)} \]

is an increasing function of \( g_s(\omega) \), so the relative underestimation will be large (i.e. \( g_s(\omega) \) will be small) when the variance of the component innovation \( V(s) \) is relatively small. Underestimation of the component will be particularly intense when the stochastic variability of the component is small, resulting in the fact that the estimator is always biased towards producing an estimator more stable than the component.\(^{186}\)

In the time domain the ratio of pseudo-spectra is replaced by the ratio of pseudo-autocovariance generating function (p-ACGF)\(^{187}\) of the stationary ARIMA model \( \theta(B)d_t = \theta_s(B)\phi_n(B)b(t) \) with variance \( V(b) = k_s \), resulting in a filter:

\[ v_s(B,F) = k_s \gamma \psi((BB,FF)), \quad [7.43] \]

\(^{186}\) Ibid.

\(^{187}\) The ACGF is well defined for the stationary time series, i.e. ACGF of \( \delta_s(B)s_t \) is \( \psi \delta_s(B)_{(\omega)} = \psi \delta_s(B)_{(\omega)} \). Obviously,

\[ \delta_s(B) \text{ and } \delta_s(B) \delta_s(B) \text{ contain differencing operators that make, respectively, } s_t \text{ and } x_t \text{ stationary. Thus, for } s, \text{ the pseudo-ACGF is calculated as: } \psi \psi((BF))_{(\omega)} = \psi((BF))_{(\omega)}V(a_s). \]
where:

\[
\gamma_s(B, F) = \varphi(B)\delta(B)\varphi(F)\delta(F) V(a) - \text{ACGF of } st;
\]

\[
\gamma(B, F) = \varphi(B)\delta(B)\varphi(F)\delta(F) V(a) - \text{ACGF of } st;
\]

\[
k_s = v_{(a)\delta(F)}.
\]

Thus, Wiener-Kolmogorov filter for seasonal component \(s_t\) is expressed as:

\[
v_s(B, F) = k_s \theta_{(B)\theta(F)\delta(B)\delta(F)\varphi(F)\delta(F)}.
\] [7.44]

The final estimator of the seasonal component is given by

\[
\hat{s}_t = n_s(B, F)x(t),
\] [7.43]

where \(n_s(B, F)\) is the ACGF of the stationary ARMA model \(\theta(B)d_t = \theta_s(B)\phi_s(b(t)\) with variance \(V(b) = k_s\).

Let \(g_s(\omega)\) denote a pseudo-spectrum. One can define \(g_n(\omega) = [gg_{n0}(\omega,\omega)] \) \(g_s(\omega) = [gg_{s0}(\omega,\omega)]\) \(g_n(\omega)\), and

\[
g_s(\omega) = [g^{\theta\phi(\omega,\omega)}] g_s(\omega) = [g^{\theta\phi(\omega,\omega)}] g_s(\omega).\]

It can be noticed that \(g_n(\omega) < g_s(\omega)\) and \(g_s(\omega) < g_s(\omega)\).

The expression: \(g_s(\omega) - [g_n(\omega) + g_s(\omega)]\) is then the cross-spectrum (in time domain it is crosscovariance function between estimators). As it is positive, the MMSE yields correlated estimators. This effect emerges since variance of estimator is smaller than the variance of component. Nevertheless, if at least one non-stationary component exists, cross-correlations estimated by TRAMO/SEATS will tend to zero as cross-covariances between estimators of the components are finite. In practice, the inconvenience caused by this property will likely be of little relevance.

The final or estimator of \(\hat{s}_t\), is obtained with a doubly infinite filter, and therefore it contains an error \(e_t\) called a final estimation error, which is equal \(e_t = s_t - \hat{s}_t\).

Preliminary estimators for the components

GÓMEZ, V., and MARAVALL, A. (2001a) explain that the properties of the estimators have been derived for the final (or historical) estimators. For a finite (long enough) realization, they can be assumed to characterize the estimators for the central observations of the series, but for periods close to the beginning of the end the filter cannot be completed and some preliminary estimator has to be used. Indeed, the historical estimator shown in [7.40] is obtained for the central periods of the series. However, when $t$ approaches $T$ (last observation), the WK fitter requires observations, which are not available yet. For this reason a preliminary estimator need to be used.

To introduce the preliminary estimators it is assumed that the series is long enough so that the estimators of the components at the central period have converged to the historical ones. Under the semi-finite realisation $[x_{-\infty}, \ldots, x_T]$, where $T$ is the last observed period, the preliminary estimator of $x_{it}$ obtained at $T$ ($T - t = k \geq 0$) can be expressed as

$$\hat{x}_{it|t+k} = v_i(B, F)x_{te|T}, \quad [7.44]$$

where $v_i(B, F)$ is the WK filter and $x_{i\cdot|T}$ is the extended series, such that $x_{i\cdot|T} = x_t$ for $t \leq T$ and $x_{i\cdot|T} = \hat{x}_{it}$ for $t > T$, where $\hat{x}_{it}$ denotes the forecast of $x_t$ obtained at period $T$.

The future $k$ values necessary to apply the filter are not yet available and they are replaced by their optimal forecasts from the ARIMA model on $x_{it}$. When $k = 0$ the preliminary estimator becomes a concurrent estimator. As the forecasts are linear functions of present and past observations of $x_t$, the preliminary estimator $\hat{x}_{it}$ obtained with the forecasts will be a truncated filter applied to $x_t$. Similarly, for $1 < t < k$ the estimator will use backcasts, and hence yield preliminary estimators for the starting period. This truncated filter will be neither centered nor symmetric. As a result, a phase effect occurs$^{189}$.

When a new observation $x_{T+1}$ becomes available the forecast $\hat{x}_{T+1|T}$ is replaced by the observation and the forecast $x_{T+j|T}$, $j > 1$ are updated to $x_{T+j|T+1}$ resulting in the estimation error$^{201}$. The total error in the preliminary estimator $d_{it|t+k}$ is expressed as a sum of the final estimation error ($e_{it}$) and the revision error ($r_{it|t+k}$), i.e.:

$$d_{it|t+k} = x_{it} - \hat{x}_{it|t+k} = (x_{it} - \hat{x}_{it}) + (\hat{x}_{it} - \hat{x}_{it|t+k}) = e_{it} + r_{it|t+k}, \quad [7.45]$$

where:

- $x_{it} - i^{th}$ component;

---


\[ \hat{x}_{i|t+k} \] the estimator of \( x_{it} \) when the last observation is \( x_{t+k} \).

Therefore the preliminary estimator is a subject of not only the final error but also a revision error, which are orthogonal to each other\(^{190}\). The revision error decreases as \( k \) increases, until it can be assumed equal to 0 for large enough \( k \).

It is worth remembering that SEATS estimates the unobservable components of a time series so the “true” components are never observed. Therefore, MARAVALL, A. (2009) stresses that the error in the historical estimator is more of academic rather than practical interest. In practice, interest centers on revisions. (...) the revision standard deviation will be an indicator of how far we can expect to be from the optimal estimator that will be eventually attained, and the speed of convergence of \( \Theta(B)^{-1} \) will dictate the speed of convergence of the preliminary estimator to the historical one. The analysis of an error is therefore useful for making decision concerning the revision policy, including the policy for revisions and horizon of revisions.

9.1.2.3. PsiE-weights

The estimator of the component is calculated as \( \hat{x}_{it} = v_s(B, F)x_t \). By replacing \( x_{it} = \gamma \delta(B) a_t \), the component estimator can be expressed as\(^{191}\):

\[ \hat{x}_{it} = \xi_s(B, F) a_t, \quad [7.46] \]

where \( \xi_s(B, F) = \cdots + \xi_1 B + \cdots + \xi_0 + \xi_{-1} F + \cdots + \xi_{-j} F^j \).

This representation shows the estimator as a filter applied to the innovation \( a_t \) rather than on the series \( x_t \).\(^{204}\) Hence, the filter from [7.46] can be divided into two components: the first one, i.e. \( \cdots + \xi_1 B + \cdots + \xi_0 \) applies to prior and concurrent innovations, the second one, i.e. \( \xi_{-1} F + \cdots + \xi_{-j} F^j \) applies to future (i.e. posterior to \( t \)) innovations. Consequently, \( \xi_t \) determines the contribution of \( a_{t-j} \) to \( \hat{s}_t \), while \( \xi_{-j} \) determines the contribution of \( a_{t+j} \) to \( \hat{s}_t \). Finally, the estimator of the component can be expressed as:

\[ \hat{x}_{it} = \xi_t(B)^{-a_t} + \xi_t(F)^{a_{t+1}}, \quad [7.47] \]

where:

\( \xi_t(B)^{-a_t} \) is an effect of starting conditions, present and past innovations in series; \( \xi_t(F)^{a_{t+1}} \) is an effect of future innovations.

---

\(^{190}\) MARAVALL, A. (2009).

\(^{191}\) The section is based on KAISER, R., and MARAVALL, A. (2000). \(^{204}\) See 7.1.2.3. For further details see MARAVALL, A. (2008).
It can be shown that $\xi, \ldots, \xi_{-1}$, and $\xi_{j}, \ldots, \xi_{0}$ are divergent and convergent. From [7.47], the concurrent estimator is equal to

$$\hat{x}_{it} = E_x x_t = E \hat{x}_t = \xi_t(B) - at$$

so that the revision

$$r_{it} = \hat{x}_{it} - \hat{x}_{it}$$

is the convergent moving average

$$r_{it} = \xi_t(F)^t a_{t+1}$$

a zero-mean stationary process. As a result, historical and preliminary estimators are cointegrated. From expression [7.50] the relative size of the full revision and the speed of convergence can be obtained.

9.1.3. X-11

A complete documentation of the X-11 method is available in LADIRAY, D., and QUENNEVILLE, B. (2001). The X-11 program is the result of a long tradition of non-parametric smoothing based on moving averages, which are weighted averages of a moving span of a time series (see hereafter). Moving averages have two important drawbacks:

- They are not resistant and might be deeply impacted by outliers;
- The smoothing of the ends of the series cannot be done except with asymmetric moving averages which introduce phase-shifts and delays in the detection of turning points.

These drawbacks adversely affect the X-11 output and stimulate the development of this method. To overcome these flaws first the series are modelled with a RegARIMA model that calculates forecasts and estimates the regression effects. Therefore, the seasonal adjustment process is divided into two parts.

- In a first step, the RegARIMA model is used to clean the series from non-linearities, mainly outliers and calendar effects. A global ARIMA model is adjusted to the series in order to compute the forecasts.
- In a second step, an enhanced version of the X-11 algorithm is used to compute the trendcycle, the seasonal component and the irregular component.
9.1.3.1. Moving averages

The moving average of coefficients \( \{ \theta \} \) is defined as:

\[
M(X_t) = \sum_{k=-p}^{+f} \theta_k X_{t+k}
\]

The value at time \( t \) of the series is therefore replaced by a weighted average of \( p \) "past" values of the series, the current value, and \( f \) "future" values of the series. The quantity \( p + f + 1 \) is called the moving average order. When \( p \) is equal to \( f \), that is, when the number of points in the past is the same as the number of points in the future, the moving average is said to be centred. If, in addition, \( \theta_{-k} = \theta_k \) for any \( k \), the moving average \( M \) is said to be symmetric. One of the simplest moving averages is the symmetric moving average of order \( P = 2p + 1 \) where all the weights are equal to \( \frac{1}{p} \).

This moving average formula works well for all time series observations, except for the first \( p \) values and last \( f \) values. Generally, with a moving average of order \( p + f + 1 \) calculated for instant \( t \) with points \( p \) in the past and points \( f \) in the future, it will be impossible to smooth out the first \( p \) values and the last \( f \) values of the series because of lack of input to the moving average formula.

In the X-11 method, symmetric moving averages play an important role as they do not introduce any phase-shift in the smoothed series. But, to avoid losing information at the series ends, they are either supplemented by \textit{ad hoc} asymmetric moving averages or applied on the series extended by forecasts.
For the estimation of the seasonal component, X-13ARIMA-SEATS uses $P \times Q$ composite moving averages, obtained by composing a simple moving average of order $P$, whose coefficients are all equal to $1$, and a simple moving average of order $Q$, whose coefficients are all equal to $1$.

In the estimation of the trend-cycle, X-13ARIMA-SEATS uses Henderson moving averages which have been chosen for their smoothing properties. The coefficients of a Henderson moving average of order $2p + 1$ may be calculated using the formula:

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

where: $n = p + 2$.

9.1.3.2. The basic algorithm of the X-11 method

The X-11 method is based on an iterative principle of estimation of the different components using appropriate moving averages at each step of the algorithm. The successive results are saved in tables. The list of the X-11 tables displayed in JDemetra+ is included at the end of this section.

The basic algorithm of the X-11 method will be presented for a monthly time series $X_t$ that is assumed to be decomposable into trend-cycle, seasonality and irregular component according to an additive model $X_t = TC_t + S_t + I_t$.

A simple seasonal adjustment algorithm can be thought of in eight steps:

1. **Estimation of Trend by $2 \times 12$ moving average:**

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

The moving average used here is a $2 \times 12$ moving average, with coefficients, $\frac{1}{24}{1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1}$ that retains linear trends, eliminates order-12 constant seasonality and minimises the variance of the irregular component.

2. **Estimation of the Seasonal-Irregular component:**

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

3. **Estimation of the Seasonal component by $3 \times 3$ moving average over each month:**
\[ S(t(1)) - M_{3\times3}[(S_t + I_t)(1)]. \]  

The moving average used here is a \(3 \times 3\) moving average over 5 terms, with coefficients \(\frac{1}{9}\{1, 2, 3, 2, 1\}\). The seasonal component is then centred using a \(2\times12\) moving average.

\[ S_t(1) = S_t(1) - M_{2\times12}(S(t(1))). \]  

4. Estimation of the seasonally adjusted series:

\[ SA(t(1)) = (TC_t + I_t)(1) = X_t - S(t(1)). \]  

This first estimation of the seasonally adjusted series must, by construction, contain less seasonality. The X-11 method again executes the algorithm presented above, changing the moving averages to take this property into account.

5. Estimation of Trend by 13-term Henderson moving average:

\[ TC_t(2) = H_{13}(SA(t(1))). \]  

Henderson moving averages, while they do not have special properties in terms of eliminating seasonality (limited or none at this stage), have a very good smoothing power and retain a local polynomial trend of degree 2 and preserve a local polynomial trend of degree 3.

6. Estimation of the Seasonal-Irregular component:

\[ (S_t + I_t)(2) = X_t - TC_t(2). \]  

7. Estimation of the Seasonal component by \(3 \times 5\) moving average over each month:

\[ S_t(2) - M_{3\times3}[(S_t + I_t)(2)]. \]  

8. The moving average used here is a \(3 \times 5\) moving average over 7 terms, of coefficients \(\frac{1}{15}\{1, 2, 3, 3, 3, 2, 1\}\) and retains linear trends. The coefficients are then normalised such that their sum over the whole 12-month period is approximately cancelled out:

\[ \tilde{S}_t(2) = S_t(2) - M_{2\times12}(S_t(2)). \]  

9. Estimation of the seasonally adjusted series:

\[ SA(t(2)) = (TC_t + I_t)(2) = X_t - S_t(2). \]
The whole difficulty lies, then, in the choice of the moving averages used for the estimation of the trend in steps 1 and 5 on the one hand, and for the estimation of the seasonal component in steps 3 and 5. The course of the algorithm in the form that is implemented in JDemetra+ is presented in Figure 9.11: A workflow diagram for the X-11 algorithm. Source: Based upon training material from the Deutsche Bundesbank. The adjustment for trading day effects, which is present in the original X-11 program, is omitted here, as since calendar correction is performed by the RegARIMA model, JDemetra+ does not perform further adjustment for these effects in the decomposition step.

9.1.3.2.1. The iterative principle of X-11

To evaluate the different components of a series, while taking into account the possible presence of extreme observations, X-11 will proceed iteratively: estimation of components, search for disruptive effects in the irregular component, estimation of components over a corrected series, search for disruptive effects in the irregular component, and so on.

The Census X-11 program presents four processing stages (A, B, C, and D), plus 3 stages, E, F, and G, that propose statistics and charts and are not part of the decomposition per se. In stages B, C and D the basic algorithm is used as is indicated in Figure 9.12: A workflow diagram for the X-11 algorithm implemented in JDemetra+. Source: Based upon training material from the Deutsche Bundesbank.
Part A: Pre-adjustments

This part, which is not obligatory, corresponds in X-13ARIMA-SEATS to the first cleaning of the series done using the RegARIMA facilities: detection and estimation of outliers and calendar effects (trading day and Easter), forecasts and backcasts\(^\text{192}\) of the series. Based on these results, the program calculates prior adjustment factors that are applied to the raw series. The series thus corrected, Table B1 of the printouts, then proceeds to part B.

Part B: First automatic correction of the series

This stage consists of a first estimation and down-weighting of the extreme observations and, if requested, a first estimation of the calendar effects. This stage is performed by applying the basic algorithm detailed earlier. These operations lead to Table B20, adjustment values for extreme observations, used to correct the unadjusted series and result in the series from Table C1.

---

\(^{192}\) This is a general estimation procedure used by the U.S. Census Bureau. JDemetra+ does not calculate backcasts for X13ARIMA-SEATS.
Part C: Second automatic correction of the series
Applying the basic algorithm once again, this part leads to a more precise estimation of replacement values of the extreme observations (Table C20). The series, finally "cleaned up", is shown in Table D1 of the printouts.

Part D: Seasonal adjustment
This part, at which our basic algorithm is applied for the last time, is that of the seasonal adjustment, as it leads to final estimates:

- of the seasonal component (Table D10);
- of the seasonally adjusted series (Table D11);
- of the trend component (Table D12);
- of the irregular component (Table D13).

Part E: Components modified for large extreme values
Parts E includes:

- Components modified for large extreme values;
- Comparison the annual totals of the raw time series and seasonally adjusted time series;
- Changes in the final seasonally adjusted series;
- Changes in the final trend-cycle;
- Robust estimation of the final seasonally adjusted series.

The results from part E are used in part F to calculate the quality measures.

Part F: Seasonal adjustment quality measures
Part F contains statistics for judging the quality of the seasonal adjustment. JDemetra+ presents selected output for part F, i.e.:

- M and Q statistics;
- Tables.

Part G: Graphics
Part G presents spectra estimated for:

- Raw time series adjusted a priori (Table B1);
• Seasonally adjusted time series modified for large extreme values (Table E2); • Final irregular component adjusted for large extreme values (Table E3).

Originally, graphics were displayed in character mode. In JDematra+, these graphics are replaced favourably by the usual graphics software.

The Henderson moving average and the trend estimation

In iteration B (Table B7), iteration C (Table C7) and iteration D (Table D7 and Table D12) the trend component is extracted from an estimate of the seasonally adjusted series using Henderson moving averages. The length of the Henderson filter is chosen automatically by X-13ARIMA-SEATS in a two-step procedure.

It is possible to specify the length of the Henderson moving average to be used. X-13ARIMASEATS provides an automatic choice between a 9-term, a 13-term or a 23-term moving average. The automatic choice of the order of the moving average is based on the value of an indicator called $\gamma$ ratio which compares the magnitude of period-on-period movements in the irregular component $I$ with those in the trend. The larger the ratio, the higher the order of the moving average selected. Moreover, X-13ARIMA-SEATS allows the user to choose manually any odd-numbered Henderson moving average. The procedure used in each part is very similar; the only differences are the number of options available and the treatment of the observations in the both ends of the series. The procedure below is applied for a monthly time series.

In order to calculate $\gamma$ ratio a first decomposition of the SA series (seasonally adjusted) is computed using a 13-term Henderson moving average.

For both the trend-cycle ($C$) and irregular ($I$) components, the average of the absolute values for monthly growth rates (multiplicative model) or for monthly growth (additive model) are computed. They are denoted as $\overline{C}$ and $\overline{I}$, receptively, where $\overline{C} = \frac{1}{n-1} \sum_{t=2}^{n} \left| C_t - C_{t-1} \right|$ and $\overline{I} = \frac{1}{n-1} \sum_{t=2}^{n} \left| I_t - I_{t-1} \right|$. Then the value of $\gamma$ ratio is checked and in iteration B:

• If the ratio is smaller than 1, a 9-term Henderson moving average is selected; • Otherwise, a 13-term Henderson moving average is selected.
Then the trend is computed by applying the selected Henderson filter to the seasonally adjusted series from Table B6. The observations at the beginning and at the end of the time series that cannot be computed by means of symmetric Henderson filters are estimated by ad hoc asymmetric moving averages.

In iterations C and D:

- If the ratio is smaller than 1, a 9-term Henderson moving average is selected; •
  - If the ratio is greater than 3.5, a 23-term Henderson moving average is selected.
- Otherwise, a 13-term Henderson moving average is selected.

The trend is computed by applying selected Henderson filter to the seasonally adjusted series from Table C6, Table D7 or Table D12, accordingly. At both ends of the series, where a central Henderson filter cannot be applied, the asymmetric ends weights for the 7 term Henderson filter are used.

9.1.3.2.2. Choosing the composite moving averages when estimating the seasonal component

In iteration D, Table D10 shows an estimate of the seasonal factors implemented on the basis of the modified SI (Seasonal-Irregular) factors estimated in Tables D4 and D9bis. This component will have to be smoothed to estimate the seasonal component; depending on the importance of the irregular in the Seasonal-Irregular component, we will have to use moving averages of varying length as in the estimate of the Trend/Cycle where the \( c \) ratio was used to select the length of the Henderson moving average. The estimation includes several steps.

**Step 1: Estimating the irregular and seasonal components**

An estimate of the seasonal component is obtained by smoothing, month by month and therefore column by column, Table D9bis using a simple 7-term moving average, i.e. of coefficients \( \frac{1}{7}(1, 1, 1, 1, 1, 1) \). In order not to lose three points at the beginning and end of each column, all columns are completed arbitrarily. Let us assume that the column that corresponds to the month is composed of \( N \) values \( \{x_1, x_2, x_3, \ldots, x_{N-1}, x_N\} \). It will be transformed into a series \( \{x_{-2}, x_{-1}, x_0, x_1, x_2, x_3, \ldots, x_{N-1}, x_Nx_{N+1}, x_{N+1}, x_{N+2}, x_{N+3}\} \) with: \( x_{-2} = x_{-1} = x_0 = x \frac{1}{3} + x + x \) and \( x_{N+1} = x_N + 2 = x_{N+2} = x_{N+3} = x^2 + x_{N-2} + x_{N-1} + x_{N-3}. \) We then have the required estimates: \( S = M_7(D9bis) \) and \( l = D9bis - S. \)
Step 2: Calculating the Moving Seasonality Ratios

For each $i^{th}$ month the mean annual changes for each component is obtained by calculating $S_i = \frac{1}{N_i-1} \sum_{t=2}^{N_i} |S_{i,t} - S_{i,t-1}|$ and $I_i = \frac{1}{N_i-1} \sum_{t=2}^{N_i} |I_{i,t} - I_{i,t-1}|$, where $N_i$ refers to the number of months $i$ in $i$, the data, and the moving seasonality ratio of month $i: MSR_i = \sum_{i} S_i$. These ratios are presented in Details of the Quality Measures node under the Decomposition (X11) section. These ratios are used to compare the year-on-year changes in the irregular component with those in the seasonal component. The idea is to obtain, for each month, an indicator capable of selecting the appropriate moving average for the removal of any noise and providing a good estimate of the seasonal factor. The higher the ratio, the more erratic the series, and the greater the order of the moving average should be used. As for the rest, by default the program selects the same moving average for each month, but the user can select different moving averages for each month.

Step 3: Calculating the overall Moving Seasonality Ratio

The overall Moving Seasonality Ratio is calculated as follows:

$$MSR_i = \frac{\sum_{i} S_i}{\sum_{i} N_i}$$

[7.63]

Step 4: Selecting a moving average and estimating the seasonal component

Depending on the value of the ratio, the program automatically selects a moving average that is applied, column by column (i.e. month by month) to the Seasonal/Irregular component in Table D8 modified, for extreme values, using values in Table D9.

The default selection procedure of a moving average is based on the Moving Seasonality Ratio in the following way (Figure 9.13):

- If this ratio occurs within zone A ($MSR < 2.5$), a $3 \times 3$ moving average is used; if it occurs within zone C ($3.5 < MSR < 5.5$), a $3 \times 5$ moving average is selected; if it occurs within zone E ($MSR > 6.5$), a $3 \times 9$ moving average is used;

- If the MSR occurs within zone B or D, one year of observations is removed from the end of the series, and the MSR is recalculated. If the ratio again occurs within zones B or D, we start over
again, removing a maximum of five years of observations. If this does not work, i.e. if we are again within zones B or D, a 3x5 moving average is selected.

The chosen symmetric moving average corresponds, as the case may be 5 (3x3), 7 (3x5) or 11 (3x9) terms, and therefore does not provide an estimate for the values of seasonal factors in the first 2 (or 3 or 5) and the last 2 (or 3 or 5) years. These are then calculated using ad hoc asymmetric moving averages.

![Figure 9.13: Moving average selection procedure, source: DAGUM, E. B. (1999).](image)

9.1.3.2.3. Identification and replacement of extreme values

X-13ARIMA-SEATS detects and removes outliers in the RegARIMA part. However, if there is a seasonal heteroscedasticity in a time series i.e. the variance of the irregular component is different in different calendar months. Examples for this effect could be the weather and snow-dependent output of the construction sector in Germany during winter, or changes in Christmas allowances in Germany and resulting from this a transformation in retail trade turnover before Christmas. The ARIMA model is not on its own able to cope with this characteristic. The practical consequence is given by the detection of additional extreme values by X-11. This may not be appropriate if the seasonal heteroscedasticity is produced by political interventions or other influences. The ARIMA models assume a constant variance and are therefore not by themselves able to cope with this problem. Choosing longer (in the case of diverging weather conditions in the winter time for the construction sector) or shorter filters (in the case of a changing pattern of retail trade turnover in the Christmas time) may be reasonable in such cases. It may even be sensible to take into account the possibility of period-specific (e.g. month-specific) standard deviations which can be done by changing the default setting of the calendarsigma parameter (see 5.1.2.2) The value of the calendarsigma parameter will have an impact on the method of calculation of the moving standard deviation in the procedure for extreme values detection presented below.

*Step 1: Estimating the seasonal component*
The seasonal component is estimated by smoothing Seasonal-Irregular component separately for
each period using (3 × 3) moving average, i.e.:
1,0,0,0,0,0,0,0,0,0,0,0,
2,0,0,0,0,0,0,0,0,0,0,0,

[7.64]

× 3,0,0,0,0,0,0,0,0,0,0,0, . 2,0,0,0,0,0,0,0,0,0,0,0,
{1,0,0,0,0,0,0,0,0,0,0,0,}
Step 2: Normalizing the seasonal factors
The preliminary seasonal factors are normalized in such a way that for one year their average is
equal to zero (additive model) or to unity (multiplicative model).
Step3: Estimating the irregular component
The initial normalized seasonal factors are removed from the Seasonal-Irregular component to
provide an estimate of the irregular component.
Step 4: Calculating a moving standard deviation
By default, a moving standard deviation of the irregular component is calculated at five-year
intervals. Each standard deviation is associated with the central year used to calculate it. The values
in the central year, which in the absolute terms deviate from average by more than the Usigma
parameter (by default, 2.5 standard deviations, see 7.1.2.2) are marked as extreme values and
assigned a zero weight. After excluding the extreme values the moving standard deviation is
calculated once again.
Step 5: Detecting extreme values and weighting the irregular
The default settings for assigning a weight to each value of irregular component are:
▪

Values which are more than Usigma (2.5, by default) standard deviations away (in the
absolute terms) from the 0 (additive) or 1 (multiplicative) are assigned a zero weight;

▪

Values which are less than 1.5 standard deviations away (in the absolute terms) from the 0
(additive) or 1 (multiplicative) are assigned a full weight (equal to one);

▪

Values which are lie between 1.5 and 2.5 standard deviations away (in the absolute terms)
from the 0 (additive) or 1 (multiplicative) are assigned a weight that varies linearly between
0 and 1 depending on their position.

250


The default boundaries for the detection of the extreme values can be changed with \textit{LSigma} and \textit{USigma} parameters (see 7.1.2.2).

\textit{Step 6: Adjusting extreme values of the seasonal-irregular component}

Values which are assigned weights that are less than 1 are regarded as extreme. Those values are replaced by a weighted average of five values:

- The value itself with its weight;
- The two preceding values, for the same period, having full weight;
- The next two values, for the same period, having full weight.

This general algorithm is used with some modification in parts B and C for detection and replacement of extreme values.

\textit{9.1.3.2.4. X-11 tables}

The list of tables produced by JDemetra+ is presented below. It is not identical to the output produced by the original X-11 program.

Part A – Preliminary Estimation of Outliers and Calendar Effects. This part includes prior modifications to the original data made in the RegARIMA part:

- Table A1 – Original series;
- Table A1a – Forecast of Original Series;
- Table A2 – Leap year effect;
- Table A6 – Trading Day effect (1 or 6 variables);
- Table A7 – The Easter effect;
- Table A8 – Total Outlier Effect;
- Table A8i – Additive outlier effect;
- Table A8t – Level shift effect;
- Table A8s – Transitory effect;
- Table A9 – Effect of user-defined regression variables assigned to the seasonally adjusted series or for which the component has not been defined;
- Table 9sa – Effect of user-defined regression variables assigned to the seasonally adjusted series;
- Table9u – Effect of user-defined regression variables for which the component has not been defined.

Part B – Preliminary Estimation of the Time Series Components:
Part C – Final Estimation of Extreme Values and Calendar Effects:

- Table C1 – Modified Raw Series;
- Table C2 – Trend (preliminary estimation using composite moving average);
- Table C4 – Modified Seasonal-Irregular Component;
- Table C5 – Seasonal Component;
- Table C6 – Seasonally Adjusted Series;
- Table C7 – Trend (estimation using Henderson moving average);
- Table C9 – Seasonal-Irregular Component;
- Table C10 – Seasonal Component;
- Table C11 – Seasonally Adjusted Series;
- Table C13 – Irregular Component;
- Table C17 – Preliminary Weights for the Irregular;
- Table B20 – Adjustment Values for Extreme Irregulars.

Part D – Final Estimation of the Different Components:

- Table D1 – Modified Raw Series;
- Table D2 – Trend (preliminary estimation using composite moving average);
- Table D4 – Modified Seasonal-Irregular Component;
- Table D5 – Seasonal Component;
- Table D6 – Seasonally Adjusted Series;
- Table D7 – Trend (estimation using Henderson moving average);
9.2. Calendar effects in JDemetra+

The following description of the calendar effects in JDemetra+ is strictly based on PALATE, J. (2014).
A natural way for modelling calendar effects consists of distributing the days of each period into different groups. The regression variable corresponding to a type of day (a group) is simply defined by the number of days it contains for each period. Usual classifications are:

- Trading days (7 groups): each day of the week defines a group (Mondays,...,Sundays);
- Working days (2 groups): week days and weekends.

The definition of a group could involve partial days. For instance, we could consider that one half of Saturdays belong to week days and the second half to weekends.

Usually, specific holidays are handled as Sundays and they are included in the group corresponding to "non-working days". This approach assumes that the economic activity on national holidays is the same (or very close to) the level of activity that is typical for Sundays. Alternatively, specific holidays can be considered separately, e.g. by the specification that divided days into three groups:

- Working days (Mondays to Fridays, except for specific holidays);
- Non-working days (Saturdays and Sundays, except for specific holidays);
- Specific holidays.

9.2.1. Mean and seasonal effects of calendar variables

The calendar effects produced by the regression variables that fulfil the definition presented above include a mean effect (i.e. an effect that is independent of the period) and a seasonal effect (i.e. an effect that is dependent of the period and on average it is equal to 0). Such an outcome is inappropriate, as in the usual decomposition of a series the mean effect should be allocated to the trend component and the fixed seasonal effect should be affected to the corresponding component. Therefore, the actual calendar effect should only contain effects that don’t belong to the other components.

In the context of JDemetra+ the mean effect and the seasonal effect are long term theoretical effects rather than the effects computed on the time span of the considered series (which should be continuously revised).

The mean effect of a calendar variable is the average number of days in its group. Taking into account that one year has on average 365.25 days, the monthly mean effects for a working days are, as shown in Table 9.2, 21.7411 for week days and 8.696 for weekends.

Table 9.2: Monthly mean effects for the Working day variable.
The number of days by period is highly seasonal, as, apart from February, the length of each month is the same every year. For this reason, any set of calendar variables will contain, at least in some variables, a significant seasonal effect, which is defined as the average number of days by period (Januarys..., first quarters...) outside the mean effect.Removing that fixed seasonal effects consists of removing for each period the long term average of days that belong to it. The calculation of a seasonal effect for the working days classification is presented in Table 9.3.

### Table 9.3: The mean effect and the seasonal effect for the calendar periods.

<table>
<thead>
<tr>
<th>Period</th>
<th>Average number of days</th>
<th>Average number of week days</th>
<th>Mean effect</th>
<th>Seasonal effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>31</td>
<td>31*5/7=22.1429</td>
<td>21.7411</td>
<td>0.4018</td>
</tr>
<tr>
<td>March</td>
<td>31</td>
<td>31*5/7=22.1429</td>
<td>21.7411</td>
<td>0.4018</td>
</tr>
<tr>
<td>April</td>
<td>30</td>
<td>30*5/7=21.4286</td>
<td>21.7411</td>
<td>-0.3125</td>
</tr>
<tr>
<td>May</td>
<td>31</td>
<td>31*5/7=22.1429</td>
<td>21.7411</td>
<td>0.4018</td>
</tr>
<tr>
<td>June</td>
<td>30</td>
<td>30*5/7=21.4286</td>
<td>21.7411</td>
<td>-0.3125</td>
</tr>
<tr>
<td>July</td>
<td>31</td>
<td>31*5/7=22.1429</td>
<td>21.7411</td>
<td>0.4018</td>
</tr>
<tr>
<td>August</td>
<td>31</td>
<td>31*5/7=22.1429</td>
<td>21.7411</td>
<td>0.4018</td>
</tr>
<tr>
<td>September</td>
<td>30</td>
<td>30*5/7=21.4286</td>
<td>21.7411</td>
<td>-0.3125</td>
</tr>
<tr>
<td>October</td>
<td>31</td>
<td>31*5/7=22.1429</td>
<td>21.7411</td>
<td>0.4018</td>
</tr>
<tr>
<td>November</td>
<td>30</td>
<td>30*5/7=21.4286</td>
<td>21.7411</td>
<td>-0.3125</td>
</tr>
<tr>
<td>December</td>
<td>31</td>
<td>31*5/7=22.1429</td>
<td>21.7411</td>
<td>0.4018</td>
</tr>
<tr>
<td>Total</td>
<td>365.25</td>
<td>260.8929</td>
<td>260.8929</td>
<td>0</td>
</tr>
</tbody>
</table>

For a given time span, the actual calendar effect for week days can be easily calculated as the difference between the number of week days in a specific period and the sum of the mean effect and the seasonal effect assigned to this period, as it is shown in Table 9.4 for the period 01.2013 - 06.2013.

### Table 9.4: The calendar effect for the period 01.2013 - 06.2013.

<table>
<thead>
<tr>
<th>Time period (t)</th>
<th>Week days</th>
<th>Mean effect</th>
<th>Seasonal effect</th>
<th>Calendar effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan-2013</td>
<td>23</td>
<td>21.7411</td>
<td>0.4018</td>
<td>0.8571</td>
</tr>
</tbody>
</table>

255
The distinction between the mean effect and the seasonal effect is usually unnecessary. Those effects can be considered together (simply called mean effects) and be computed by removing from each calendar variable its average number of days by period. These global means effect are considered in the next section.

9.2.2. Linear transformations of the calendar variables

As far as the RegARIMA and the TRAMO models are considered, any non-degenerated linear transformation of the calendar variables can be used. It will produce the same results (likelihood, residuals, parameters, joint effect of the calendar variables, joint F-test on the coefficients of the calendar variables,...). The linearised series that will be further decomposed is invariant to any linear transformation of the calendar variables.

However, it should be mentioned that choices of calendar corrections based on the tests on the individual t statistics are dependent on the transformation, which is rather arbitrary. This is the case in old versions of TRAMO/SEATS. That is why the joint F-test (as in the version of TRAMO/SEATS implemented in TSW+) should be preferred.

An example of a linear transformation is the calculation of the contrast variables. In the case of the usual trading day variables, they are defined by the following transformation: the 6 contrast variables (No. (Mondays) – No. (Sundays),… No. (Saturdays) – No. (Sundays)) used with the length of period.

\[
\begin{bmatrix}
1 \\
0 \\
0 \\
0 \\
0 \\
0 \\
\end{bmatrix}
-1 Mon \\
-1 Tue \\
\begin{bmatrix}
1 \\
1 \\
1 \\
1 \\
1 \\
1 \\
\end{bmatrix}
\begin{bmatrix}
Mon – Sun \\
Tue – Sun \\
Wed – Sun \\
Thu – Sun \\
Fri – Sun \\
Sat – Sun \\
\end{bmatrix}
\]

For the usual working day variables, two variables are used: one contrast variable and the length of period.
\[1 - 5/2 \] \(Week\) = \([Length\, Contrast\, of\, period\, week\,].\)  

\[7.66\]

1 1 \(Weekend\)

The \textit{Length of period} variable is defined as a deviation from the length of the month (in days) and the average month length, which is equal to 30.4375. Instead, the leap-year variable can be used here (see 4.1.1.3 or 4.1.2.3)\textsuperscript{193}.

Such transformations have several advantages. They suppress from the contrast variables the mean and the seasonal effects, which are concentrated in the last variable. So, they lead to fewer correlated variables, which are more appropriate to be included in the regression model. The sum of the effects of each day of the week estimated with the trading (working) day contrast variables cancel out.

\textbf{9.2.3. Handling of specific holidays}

Three types of holidays are implemented in JDemetra+:

- Fixed days, corresponding to the fixed dates in the year (e.g. New Year, Christmas).
- Easter related days, corresponding to the days that are defined in relation to Easter (e.g. Easter +/- n days; example: Ascension, Pentecost).
- Fixed week days, corresponding to the fixed days in a given week of a given month (e.g. Labor Day celebrated in the USA on the first Monday of September).

From a conceptual point of view, specific holidays are handled in exactly the same way as the other days. It should be decided, however, to which group of days they belong. Usually they are handled as Sundays. This convention is also used in JDemetra+. Therefore, except if the holiday falls on a Sunday, the appearance of a holiday leads to correction in two groups, i.e. in the group that contains the weekday, in which holiday falls, and the group that contains the Sundays.

Country specific holidays have an impact on the mean and the seasonal effects of calendar effects. Therefore, the appropriate corrections to the number of particular days (which are usually the basis for the definition of other calendar variables) should be applied, following the kind of holidays. These corrections are applied to the period(s) that may contain the holiday. The long term corrections in JDemetra+ don’t take into account the fact that some moving holidays could fall on the same day (for instance the May Day and the Ascension). However, those events are exceptional, and their impact on the final result is usually not significant.

\textsuperscript{193} GÓMEZ, V., and MARAVALL, A (2001b).
9.2.3.1. Fixed day

The probability that the holiday falls on a given day of the week is 1/7. Therefore, the probability to have 1 day more that is treated like Sunday is 6/7. The effect on the means for the period that contains the fixed day is presented in Table 9.5 (the correction on the calendar effect has the opposite sign).

Table 9.5: The effect of the fixed holiday on the period, in which it occurred.

<table>
<thead>
<tr>
<th>Sundays</th>
<th>Others days</th>
<th>Contrast variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ 6/7</td>
<td>- 1/7</td>
<td>- 1/7 - (+ 6/7)= -1</td>
</tr>
</tbody>
</table>

9.2.3.2. Easter related days

Easter related days always fall the same week day (denoted as X in Table 9.6). However, they can fall during different periods (months or quarters). Suppose that, taking into account the distribution of the dates for Easter and the fact that this holiday falls in one of two periods, the probability that Easter falls during the period \( m \) is \( p \), which implies that the probability that it falls in the period \( m + 1 \) is \( 1 - p \). The effects of Easter on the seasonal means are presented in Table 9.6.

Table 9.6: The effects of the Easter Sunday on the seasonal means.

<table>
<thead>
<tr>
<th>Period</th>
<th>Sundays</th>
<th>Days X</th>
<th>Others days</th>
<th>Contrast X</th>
<th>Other contrasts</th>
</tr>
</thead>
<tbody>
<tr>
<td>( m )</td>
<td>+ p</td>
<td>- p</td>
<td>0</td>
<td>- 2p</td>
<td>- p</td>
</tr>
<tr>
<td>( m+1 )</td>
<td>+ (1-p)</td>
<td>- (1-p)</td>
<td>0</td>
<td>- 2×(1-p)</td>
<td>- (1-p)</td>
</tr>
</tbody>
</table>

The distribution of the dates for Easter may be approximated in different ways. One of the solutions consists of using some well-known algorithms for computing Easter on a very long period. JDemetra+ provides the Meeus/Jones/Butcher’s and the Ron Mallen’s algorithms (they are identical till year 4100, but they slightly differ after that date). Another approach consists in deriving a raw theoretical distribution based on the definition of Easter. It is the solution used for Easter related days. It is shortly explained below.

The date of Easter in the given year is the first Sunday after the full moon (the Paschal Full Moon) following the northern hemisphere’s vernal equinox. The definition is influenced by the Christian tradition, according to which the equinox is reckoned to be on 21 March\(^{194} \) and the full moon is not necessarily the astronomically correct date. However, when the full moon falls on Sunday, then Easter is delayed by one week. With this definition, the date of Easter Sunday varies between 22 March and 25 April. Taking into account that an average lunar month is 29.530595 days the approximated distribution of Easter can be derived. These calculations do not take into account the actual ecclesiastical moon calendar.

\(^{194} \) In fact, astronomical observations show that the equinox occurs on 20 March in most years.
For example, the probability that Easter Sunday falls on 25 March is 0.004838 and results from the facts that the probability that 25 March falls on a Sunday is $1/7$ and the probability that the full moon is on 21 March, 22 March, 23 March or 24 March is $5/29.53059$. The probability that Easter falls on 24 April is 0.01708 and results from the fact that the probability that 24 April is Sunday is $1/7$ and takes into account that 18 April is the last acceptable date for the full moon. Therefore the probability that the full moon is on 16 April or 17 April is $1/29.53059$ and the probability that the full moon is on 18 April is $1.53059/29.53059$.

### Table 9.7: The approximated distribution of Easter dates.

<table>
<thead>
<tr>
<th>Day</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>22 March</td>
<td>$1/7 \times 1/29.53059$</td>
</tr>
<tr>
<td>23 March</td>
<td>$1/7 \times 2/29.53059$</td>
</tr>
<tr>
<td>24 March</td>
<td>$1/7 \times 3/29.53059$</td>
</tr>
<tr>
<td>25 March</td>
<td>$1/7 \times 4/29.53059$</td>
</tr>
<tr>
<td>26 March</td>
<td>$1/7 \times 5/29.53059$</td>
</tr>
<tr>
<td>27 March</td>
<td>$1/7 \times 6/29.53059$</td>
</tr>
<tr>
<td>28 March</td>
<td>$1/29.53059$</td>
</tr>
<tr>
<td>29 March</td>
<td>$1/29.53059$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>18 April</td>
<td>$1/29.53059$</td>
</tr>
<tr>
<td>19 April</td>
<td>$1/7 \times (6 + 1.53059)/29.53059$</td>
</tr>
<tr>
<td>20 April</td>
<td>$1/7 \times (5 + 1.53059)/29.53059$</td>
</tr>
<tr>
<td>21 April</td>
<td>$1/7 \times (4 + 1.53059)/29.53059$</td>
</tr>
<tr>
<td>22 April</td>
<td>$1/7 \times (3 + 1.53059)/29.53059$</td>
</tr>
<tr>
<td>23 April</td>
<td>$1/7 \times (2 + 1.53059)/29.53059$</td>
</tr>
<tr>
<td>24 April</td>
<td>$1/7 \times (1 + 1.53059)/29.53059$</td>
</tr>
<tr>
<td>25 April</td>
<td>$1/7 \times 1.53059/29.53059$</td>
</tr>
</tbody>
</table>

#### 9.2.3.3. Fixed week days

Fixed week days always fall on the same week day (denoted as X in Table 9.8) and in the same period. Their effect on the seasonal means is presented in Table 9.8.

### Table 9.8: The effect of the fixed week holiday on the period, in which it occurred.

<table>
<thead>
<tr>
<th>Sundays</th>
<th>Day X</th>
<th>Others days</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The impact of fixed week days on the regression variables is zero because the effect itself is compensated by the correction for the mean effect.

9.2.4. Impact of the mean effects on the decomposition

When the ARIMA model contains a seasonal difference - something that should always happen with calendar variables - the mean effects contained in the calendar variables are automatically eliminated, so that they don't modify the estimation. The model is indeed estimated on the series/regression variables after differencing. However, they lead to a different linearised series \( y_{lin} \). The impact of other corrections (mean and/or fixed seasonal) on the decomposition is presented in the next paragraph. Such corrections could be obtained, for instance, by applying other solutions for the long term corrections or by computing them on the time span of the series.

Now the model with "correct" calendar effects (denoted as \( C \)), i.e. effects without mean and fixed seasonal effects, can be considered. To simplify the problem, the model has no other regression effects.

For such a model the following relations hold:

\[
\begin{align*}
y_{lin} &= y - C \\
T &= F_T(y_{lin}) \\
S &= F_S(y_{lin}) + C \\
I &= F_I(y_{lin})
\end{align*}
\]

where:
T – the trend;
S – the seasonal component;
I – the irregular component;
\( F_X \) – the linear filter for the component X.

Consider next other calendar effects (\( C \)) that contain some mean (\( cm \), integrated to the final trend) and fixed seasonal effects (\( cs \), integrated to the final seasonal). The modified equations are now:

\[
\begin{align*}
C &= C + cm + cs \\
\hat{y}_{lin} &= y - C = y_{lin} - cm - cs
\end{align*}
\]
\[ T = F_T(\tilde{y}_{lin}) + cm \]  \hspace{1cm} [7.73]

\[ S = F_S(\tilde{y}_{lin}) + C + cs \]  \hspace{1cm} [7.74]  \hspace{1cm} \tilde{I} = F_I(\tilde{y}_{lin}) \]  \hspace{1cm} [7.75]

Taking into account that \( F_X \) is a linear transformation and that\(^\text{195}\)

\[ F_T(cm) = cm \]  \hspace{1cm} [7.76]

\[ F_T(cs) = 0 \]  \hspace{1cm} [7.77]

\[ F_S(cm) = 0 \]  \hspace{1cm} [7.78]

\[ F_S(cs) = cs \]  \hspace{1cm} [7.79]

\[ F_I(cm) = 0 \]  \hspace{1cm} [7.80]

\[ F_I(cs) = 0 \]  \hspace{1cm} [7.81]

The following relationships hold:

\[ T = F_T(\tilde{y}_{lin}) + cm = F_T(y_{lin}) - cm + cm = T \]  \hspace{1cm} [7.82]

\[ S = F_S(\tilde{y}_{lin}) + C + cs = F_S(y_{lin}) - cs + C + cs = S \]  \hspace{1cm} [7.83]

\[ \tilde{I} = I \]  \hspace{1cm} [7.84]

If we don’t take into account the effects and apply the same approach as in the “correct” calendar effects, we will get:

\[ T = F_T(\tilde{y}_{lin}) = T - cm \]  \hspace{1cm} [7.85]

\[ S = F_S(\tilde{y}_{lin}) + C = S + cm \]  \hspace{1cm} [7.86]

\[ I = F_I(\tilde{y}_{lin}) = I \]  \hspace{1cm} [7.87]

The trend, seasonal and seasonally adjusted series will only differ by a (usually small) constant.

\(^{195}\) In case of SEATS, the properties can be trivially derived from the matrix formulation of signal extraction. They are also valid for X-11 (additive).
In summary, the decomposition does not depend on the mean and fixed seasonal effects used for the calendar effects, provided that those effects are integrated in the corresponding final components. If these corrections are not taken into account, the main series of the decomposition will only differ by a constant.

9.2.5. Holidays with a validity period

When a holiday is valid only for a given time span, JDemetra+ applies the long term mean corrections only on the corresponding period. However, those corrections are computed in the same way as in the general case.

It is important to note that using or not using mean corrections will impact in the estimation of the RegARIMA and TRAMO models. Indeed, the mean corrections do not disappear after differencing. The differences between the SA series computed with or without mean corrections will no longer be constant.

9.2.6. Summary of the method used in JDemetra+ to compute trading day and working day effects

The computation of trading day and working days effects is performed in four steps:

1. Computation of the number of each weekday performed for all periods.
2. Calculation of the usual contrast variables for trading day and working day.
3. Correction of the contrast variables with specific holidays (for each holiday add +1 to the number of Sundays and subtract 1 from the number of days of the holiday). The correction is not performed if the holiday falls on a Sunday, taking into account the validity period of the holiday.
4. Correction of the constant variables for long term mean effects, taking into account the validity period of the holiday; see below for the different cases.

The corrections of the constant variables may receive a weight corresponding to the part of the holiday considered as a Sunday.

An example below illustrates the application of the above algorithm for the hypothetical country in which three holidays are celebrated:

- New Year (a fixed holiday, celebrated on 01 January);
- Shrove Tuesday (a moving holiday, which falls 47 days before Easter Sunday, celebrated until the end of 2012);
- Freedom day (a fixed holiday, celebrated on 25 April).
The consecutive steps in calculation of the calendar for 2012 and 2013 years are explained below. First, the number of each day of the week in the given month is calculated (Table 9.9).

**Table 9.9: Number of each weekday in different months.**

<table>
<thead>
<tr>
<th>Month</th>
<th>Mon</th>
<th>Tue</th>
<th>Wed</th>
<th>Thu</th>
<th>Fri</th>
<th>Sat</th>
<th>Sun</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan-12</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Feb-12</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Mar-12</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Apr-12</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>May-12</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Jun-12</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Jul-12</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Aug-12</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Sep-12</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Oct-12</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Nov-12</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Dec-12</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Jan-13</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Feb-13</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Mar-13</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Apr-13</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>May-13</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Jun-13</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Jul-13</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Aug-13</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Sep-13</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Oct-13</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Nov-13</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Dec-13</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

Next, the contrast variables are calculated (Table 9.10) as a result of the linear transformation applied to the variables presented in Table 9.9.
Table 9.10: Contrast variables (series corrected for leap year effects).

<table>
<thead>
<tr>
<th>Month</th>
<th>Mon</th>
<th>Tue</th>
<th>Wed</th>
<th>Thu</th>
<th>Fri</th>
<th>Sat</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan-12</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>Feb-12</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.75</td>
</tr>
<tr>
<td>Mar-12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Apr-12</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>May-12</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jun-12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Jul-12</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>Aug-12</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sep-12</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Oct-12</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Nov-12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Dec-12</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jan-13</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Feb-13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-0.25</td>
</tr>
<tr>
<td>Mar-13</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Apr-13</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>May-13</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jun-13</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jul-13</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Aug-13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Sep-13</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Oct-13</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Nov-13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Dec-13</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>0</td>
</tr>
</tbody>
</table>

In the next step the corrections for holidays is done in the following way:

- New Year: In 2012 it falls on a Sunday. Therefore no correction is applied. In 2013 it falls on a Tuesday. Consequently, the following corrections are applied to the number of each weekday in January: Tuesday -1, Sunday +1, so the following corrections are applied to the contrast variables: -2 for Tuesday and -1 for the other contrast variables.

- Shrove Tuesday: It is a fixed day of the week holiday that always falls on Tuesday. For this reason in 2012 the following corrections are applied to the number of each weekday in
February: Tuesday -1, Sunday +1, so the following corrections are applied to the contrast variables: -2 for the contrast variable associated with Tuesday, and -1 for the other contrast variables. The holiday expires at the end of 2012. Therefore no corrections are made for 2013.

- Freedom Day: In 2012 it falls on a Wednesday. Consequently, the following corrections are applied to the number of each weekday in April: Wednesday -1, Sunday +1, so the following corrections are applied to the contrast variables: -2 for Wednesday and -1 for the other contrast variables. In 2013 it falls on Thursday. Therefore, the following corrections are applied to the number of each weekday in April: Thursday -1, Sunday +1, so the following corrections are applied to the contrast variables: -2 for Thursday, and -1 for the other contrast variables.

The result of these corrections is presented in Table 9.11.

**Table 9.11: Contrast variables corrected for holidays.**

<table>
<thead>
<tr>
<th>Month</th>
<th>Mon</th>
<th>Tue</th>
<th>Wed</th>
<th>Thu</th>
<th>Fri</th>
<th>Sat</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan-12</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>Feb-12</td>
<td>-1</td>
<td>-2</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0.75</td>
</tr>
<tr>
<td>Mar-12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Apr-12</td>
<td>-1</td>
<td>-2</td>
<td>-3</td>
<td>-2</td>
<td>-2</td>
<td>-2</td>
<td>0</td>
</tr>
<tr>
<td>May-12</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jun-12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Jul-12</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>Aug-12</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sep-12</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Oct-12</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Nov-12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Dec-12</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jan-13</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>Feb-13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-0.25</td>
</tr>
<tr>
<td>Mar-13</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Apr-13</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>-2</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>May-13</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jun-13</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jul-13</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Aug-13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
Finally, the long term corrections are applied on each year of the validity period of the holiday.

- New Year: Correction on the contrasts: +1, to be applied to January of 2012 and 2013.
- Shrove Tuesday: It may fall either in February or in March. It will fall in March if Easter is on or after 17 April. Taking into account the theoretical distribution of Easter, it gives: $\text{prob(March)} = +0.22147$, $\text{prob(February)} = +0.77853$. The correction of the contrasts will be $+1.55707$ for Tuesday in February 2012 and $+0.77853$ for the other contrast variables. The correction of the contrasts will be $+0.44293$ for Tuesday in March 2012, $+0.22147$ for the other contrast variables.
- Freedom Day: Correction on the contrasts: +1, to be applied to April of 2012 and 2013.

The modifications due to the corrections described above are presented in Table 9.12.

Table 9.12: Trading day variables corrected for the long term effects.

<table>
<thead>
<tr>
<th>Month</th>
<th>Mon</th>
<th>Tue</th>
<th>Wed</th>
<th>Thu</th>
<th>Fri</th>
<th>Sat</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan-12</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Feb-12</td>
<td>-0.22115</td>
<td>-0.44229</td>
<td>0.778853</td>
<td>-0.22115</td>
<td>-0.22115</td>
<td>-0.22115</td>
<td>0.75</td>
</tr>
<tr>
<td>Mar-12</td>
<td>0.221147</td>
<td>0.442293</td>
<td>0.221147</td>
<td>1.221147</td>
<td>1.221147</td>
<td>1.221147</td>
<td>0</td>
</tr>
<tr>
<td>Apr-12</td>
<td>0</td>
<td>-1</td>
<td>-2</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>May-12</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jun-12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Jul-12</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>Aug-12</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sep-12</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Oct-12</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Nov-12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Dec-12</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jan-13</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Feb-13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-0.25</td>
</tr>
<tr>
<td>Mar-13</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
9.3. Spectral analysis

A time series $x_t$ with stationary covariance, mean $\mu$ and $k^{th}$ autocovariance $E((x_t - \mu)(x_{t-k} - \mu)) = \gamma(k)$ can be described as a weighted sum of periodic trigonometric functions: $\sin(\omega t)$ and $\cos(\omega t)$, where $\omega$ denotes frequency. Spectral analysis investigates this frequency domain representation of $x_t$ to determine how important cycles of different frequencies are in accounting for the behaviour of $x_t$.

Assuming that the autocovariances $\gamma(k)$ are absolutely summable ($\sum_{k=-\infty}^{\infty} |\gamma(k)| < \infty$), the autocovariance generating function, which summarises these autocovariances through a scalar valued function, is given by equation 7.81.\(^{196}\)

\[
(z) = \sum_{k=-\infty}^{\infty} \text{acgf} z^k \gamma(k) \tag{7.88}
\]

where $z$ denotes complex scalar.

Once the equation [7.88] is divided by $\pi$ and evaluated at some $z = e^{-i\omega} = \cos \omega - i \sin \omega$, where $-1$ and $\omega$ is a real scalar, $-\infty < \omega < \infty$, the result of this transformation is called a population spectrum $f(\omega)$ for $x_t$, given in equation 7.82.\(^{197}\)

\[
(\omega) = \frac{1}{\pi} \sum_{k=-\infty}^{\infty} f e^{-ik\omega} \gamma(k) \tag{7.89}
\]

\(^{196}\) HAMILTON, J.D. (1994).

\(^{197}\) Ibid.
Therefore, the analysis of the population spectrum in the frequency domain is equivalent to the examination of the autocovariance function in the time domain analysis; however it provides an alternative way of inspecting the process. Because \( f(\omega) \, d\omega \) is interpreted as a contribution to the variance of components with frequencies in the range \((\omega, \omega + d\omega)\), a peak in the spectrum indicates an important contribution to the variance at frequencies near the value that corresponds to this peak.

As \( e^{-i\omega} = \cos \omega - i \sin \omega \), the spectrum can be also expressed as in equation 7.90.

\[
(\omega) = \frac{1}{\pi} \sum_{k=-\infty}^{\infty} f(\cos \omega k - i \sin \omega k) \gamma(k) \tag{7.90}
\]

Since \( \gamma(k) = \gamma(-k) \) (i.e. \( \gamma(k) \) is an even function of \( k \)) and \( \sin(-x) = -\sin x \), [7.90] can be presented as equation 7.91.\(^{198}\)

\[
(\omega) = \frac{1}{\pi} \left[ \gamma(0) + 2 \sum_{k=1}^{\infty} \gamma(k) \cos \omega k \right] f \tag{7.91}
\]

which implies that if autocovariances are absolutely summable the population spectrum exists and is a continuous, real-valued function of \( \omega \). Due to the properties of trigonometric functions \((\cos(-\omega k) = \cos(\omega k) \) and \( \cos(\omega + 2\pi j) k = \cos(\omega k) \)) the spectrum is a periodic, even function of \( \omega \), symmetric around \( \omega = 0 \). Therefore, the analysis of the spectrum can be reduced to the interval \((-\pi, \pi]\). The spectrum is nonnegative for all \( \omega \in (-\pi, \pi] \).

The shortest cycle that can be distinguished in a time series lasts two periods. The frequency which corresponds to this cycle is \( \omega = \pi \) and is called the Nyquist frequency. The frequency of the longest \( 2\pi \) cycles that can be observed in the time series with \( n \) observations is \( \omega = -\frac{\pi}{n} \) and is called the fundamental (Fourier) frequency.

Note that if \( \{x_t\} \) is a white noise process with zero mean and variance \( \sigma^2 \), then for all \( |k| > 0 \) \( \gamma(k) = 0 \) and spectrum of \( \{x_t\} \) is constant \( (f(\omega) = -\sigma) \) since each frequency in the spectrum contributes equally to the variance of the process.\(^{199}\)


The aim of spectral analysis is to determine how important cycles of different frequencies are in accounting for the behaviour of a time series\textsuperscript{200}. Since spectral analysis can be used to detect the presence of periodic components, it is a natural diagnostic tool for detecting trading day effects as well as seasonal effects.\textsuperscript{201} Among the tools used for spectral analysis are the autoregressive spectrum and the periodogram. The explanations given in the subsections below derive mainly from DE ANTONIO, D., and PALATE, J. (2015) and BROCKWELL, P.J., and DAVIS, R.A. (2006).

9.3.1. Periodogram

For any given frequency $\omega$ the sample periodogram is the sample analog of the sample spectrum. In general, the periodogram is used to identify the periodic components of unknown frequency in the time series. X-13ARIMA-SEATS and TRAMO/SEATS use this tool for detecting seasonality in raw time series and seasonally adjusted series. Apart from this it is applied for checking randomness of the residuals from the ARIMA model.

To define the periodogram, first consider the vector of complex numbers\textsuperscript{202}:

\textsuperscript{200} HAMILTON, J.D. (1994).
\textsuperscript{201} SOKUP, R.J., and FINDLEY, D. F. (1999).
\textsuperscript{202} BROCKWELL, P.J., and DAVIS, R.A. (2002).
\[ x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{C} \quad [7.92] \]

where \( \mathbb{C}^n \) is the set of all column vectors with complex-valued components.

The Fourier frequencies associated with the sample size \( n \) are defined as a set of values \( \omega_j = \frac{2\pi}{n} j \), \( j = -\left\lceil \frac{n-1}{2} \right\rceil, \ldots, \left\lceil \frac{n}{2} \right\rceil \), \( -\pi < \omega_j \leq \pi \), \( j \in \mathbb{F}_n \), where \( \left\lceil h \right\rceil \) denotes the largest integer less than or equal to \( h \). The Fourier frequencies, which are called harmonics, are given by integer multiples of the fundamental frequency \( \frac{2\pi}{n} \).

Now the vectors \( e_j = n^{-1/2} (e^{-i\omega_0}, e^{-i(2\omega_0)}, \ldots, e^{-i(n\omega_0)})' \) can be defined. Vectors \( e_1, \ldots, e_n \) are orthonormal in the sense that:

\[ e_j'e_k = \begin{cases} 1, & \text{if } j = k \\ -\frac{1}{n}, & \text{if } j \neq k \end{cases} \quad [7.93] \]

where \( e_j' \) denotes the row vector whose \( k \)th component is the complex conjugate of the \( k \)th component of \( e_j \). These vectors are a basis of \( \mathbb{F}_n \) so that any \( x \in \mathbb{C}^n \) can be expressed as a sum of \( n \) components:

\[ x = \sum_{j=-\left\lceil \frac{n-1}{2} \right\rceil}^{\left\lceil \frac{n}{2} \right\rceil} a_j e_j \quad [7.94] \]

---

where the coefficients \( a_j = \mathbf{e}_j^T \mathbf{x} = n^{-1} \sum_{t=1}^{n} x_t e^{-it\omega_j} \) are derived from [7.94] by multiplying the equation on the left by \( \mathbf{e}_j \) and using [7.92].

The sequence of \( \{a_j, j \in \mathbb{Z}_n\} \) is referred as a discrete Fourier transform of \( \mathbf{x} \in \mathbb{C}^n \) and the periodogram \( I(\omega_j) \) of \( \mathbf{x} \) at Fourier frequency \( \omega_j = \frac{-\pi}{n} \) is defined as the square of the Fourier transform \( \{a_j\} \) of \( \mathbf{x} \):

\[
I(\omega_j) = |a_j|^2
\]

From [7.93] and [7.94] it can be shown that in fact the periodogram decomposes the total sum of squares \( \sum_{t=1}^{n} |x_t|^2 \) into a sums of components associated with the Fourier frequencies \( \omega_j \):

\[
\sum_{t=1}^{n} |x_t|^2 = \sum_{j=-\left[\frac{n-1}{2}\right]}^{\left[\frac{n-1}{2}\right]} I(\omega_j) |a_j|^2 \quad [7.96]
\]

If \( \mathbf{x} \in \mathbb{R}^n \), \( \omega_j \) and \( -\omega_j \) are both in \( (-\pi, -\pi] \) and \( a_j \) is presented in its polar form (i.e. \( a_j = r_j e^{i\theta_j} \)), where \( r_j \) is the modulus of \( a_j \), then [7.94] can be rewritten in the form:

\[
\mathbf{x} a_n/2 \mathbf{e}_{n/2}^T
\]

The orthonormal basis for \( \mathbb{R}^n \) is \( \{\mathbf{e}_0, \mathbf{c}_1, \mathbf{s}_1, \ldots, \mathbf{c}_{\left[\frac{n-1}{2}\right]}, \mathbf{s}_{\left[\frac{n-1}{2}\right]}, \mathbf{e}_{n/2} \text{ (excluded if n is odd)} \} \), where:

\( \mathbf{e}_0 \) is a vector composed of \( n \) elements equal to \( n^{-1/2} \), which implies that \( a_0 \mathbf{e}_0 = (n^{-1} \sum_{t=1}^{n} x_t, \ldots, n^{-1} \sum_{t=1}^{n} x_t)^T \);

\( \mathbf{c}_j = \left( \frac{1}{n^{1/2}}, \ldots, \frac{1}{n^{1/2}}, \cos \omega_j, \cos 2\omega_j, \ldots, \cos n\omega_j \right)^T \), for \( 1 \leq j \leq \left[\frac{n-1}{2}\right] \);

\( \mathbf{s}_j = \left( \frac{1}{n^{1/2}}, \ldots, \frac{1}{n^{1/2}}, \sin \omega_j, \sin 2\omega_j, \ldots, \sin n\omega_j \right)^T \), for \( 1 \leq j \leq \left[\frac{n-1}{2}\right] \);

\( \mathbf{e}_{n/2} = (-n^{1/2}, n^{1/2}, \ldots, -n^{1/2}, n^{1/2}) \).
Equation [7.97] can be seen as an OLS regression of \( x_t \) on a constant and the trigonometric terms. As the vector of explanatory variables includes \( n \) elements, the number of explanatory variables in [7.97] is equal to the number of observations. HAMILTON, J.D. (1994) shows that the explanatory variables are linearly independent, which implies that an OLS regression yields a perfect fit (i.e. without an error term). The coefficients have the form of a simple OLS projection of the data on the orthonormal basis:

\[
\hat{d}_x = \sum_{t=1}^{n} \frac{1}{\sqrt{n}} (-1)^t x_t
\]

(7.99)

\[
\hat{a}_n = \text{only when } n \text{ is even}
\]

\[
\hat{d} = \sum_{t=1}^{n} \frac{1}{\sqrt{n}} (-1)^t x_t
\]

(7.100)

\[
\hat{a}_j = \left( \frac{n}{2} \right)^{-1/2} \sum_{t=1}^{n} \left( \frac{2\pi j}{n} \right), j = 1, ..., \left( \frac{n-1}{2} \right) \ 2^{1/2} r_j \cos \theta \cos \omega_j
\]

(7.101)

\[
\beta_j = 2^{1/2} r_j \sin \theta_j = \left( \frac{n}{2} \right) \sum_{n=1}^{\left( \frac{n-1}{2} \right)} \sin \left( t \frac{\pi j}{n} \right), j = 1, ..., \left[ \frac{n-1}{2} \right]
\]

(7.102)

With [7.97] the total sum of squares \( \sum_{t=1}^{n} |x_t|^2 \) can be decomposed into \( 2 \times \left[ \frac{n-1}{2} \right] \) components corresponding to \( c_j \) and \( s_j \), which are grouped to produce the “frequency \( \omega_j \)” component for \( 1 \leq j \leq \left[ \frac{n-1}{2} \right] \). As it is shown in the Table 1.1, the value of the periodogram at the frequency \( \omega_j \) is the contribution of the \( j \)th harmonic to the total sum of squares \( \sum_{t=1}^{n} |x_t|^2 \).

Table 9.13: Decomposition of sum of squares into components corresponding to the harmonics.
Obviously, if series were random then each component $I(\omega_j)$ would have the same expectation. On the contrary, when the series contains a systematic sine component having a frequency $j$ and amplitude $A$ then the sum of squares $I(\omega_j)$ increases with $A$. In practice, it is unlikely that the frequency $j$ of an unknown systematic sine component would exactly match any of the frequencies, for which periodogram have been calculated. Therefore, the periodogram would show an increase in intensities in the immediate vicinity of $j$.\textsuperscript{204}

Note that in JDemetra+ the periodogram object corresponds exactly to the contribution to the sum of squares of the standardised data, since the series are divided by their standard deviation for computational reasons.

Using the decomposition presented in Table 1.1 the periodogram can be expressed as:

$$I(\omega_j) = r_j^2 = \frac{1}{2}(\alpha_j^2 + \beta_j^2) = \frac{1}{n} \left( \sum_{n=1}^{n} x_t \cos (\frac{2\pi j}{2} t) \right)^2 + n-1 \left( \sum_{n=1}^{n} x_t \sin (\frac{2\pi j}{2} t) \right)^2,$$

[7.103]

where $j = 0, ..., [\frac{n}{2}].$

Since $x - \bar{x}$ are generated by an orthonormal basis, and $\bar{x} = a_0 e_0$ [7.96] can be rearranged to show that the sum of squares is equal to the sum of the squared coefficients:

$$x - a_0 e_0 = \sum_{j=1}^{(n-1)/2} (\alpha_j c_j + \beta_j s_j) + a_{n/2} e_{n/2}.$$  

[7.104]

Thus the sample variance of $x_t$ can be expressed as:

$$\frac{n}{[n(n-1)/2]}$$  

[7.105]

\[ n-1 \sum (x_t - \bar{x})^2 = n-1 (\sum_{k=1}^{2} r_j^2 + 2a_{n/2}) \]

where \(a_{n/2}\) is excluded if \(n\) is odd.

The term \(2 r_j^2\) in [7.105] is then the contribution of the \(j^{th}\) harmonic to the variance and [7.105] shows then how the total variance is partitioned.

The periodogram ordinate \(I(\omega_i)\) and the autocovariance coefficient \(\gamma(k)\) are both quadratic forms of \(\{x_t\}\). It can be shown that the periodogram and autocovariance function are related and the periodogram can be written in terms of the sample autocovariance function for any non-zero Fourier frequency \(\omega_i\):\(^{205}\)

\[ I(\omega_i) = \sum_{|k| < n} \hat{\gamma}(k) e^{-ik\omega_i} = \hat{\gamma}(0) + 2 \sum_{k=1}^{\infty} \hat{\gamma}(k) \cos(k\omega_i) \]

and for the zero frequency \(I(0) = n|\bar{x}|^2\).

Once comparing [7.106] with an expression for the spectral density of a stationary process:

\[ f(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma(k) e^{-ik\omega} = \frac{1}{2\pi} [\gamma(0) + 2 \sum_{k=1}^{\infty} \gamma(k) \cos(k\omega_i)] \]

it can be noticed that the periodogram is a sample analog of the population spectrum. In fact, it can be shown that the periodogram is asymptotically unbiased but inconsistent estimator of the population spectrum \(f(\omega)\).\(^{206}\) Therefore, the periodogram is a wildly fluctuating, with high variance, estimate of the spectrum. The consistent estimator can be achieved by applying the different linear smoothing filters to the periodogram, called lag-window estimators. The lag-window estimators implemented in JDemetra+ includes square, Welch, Tukey, Barlett, Hanning and Parzen. They are described in DE ANTONIO, D., and PALATE, J. (2015). Alternatively, the

\(^{205}\) The proof is given in BROCKWELL, P.J., and DAVIS, R.A. (2006).

\(^{206}\) Ibid.
model-based consistent estimation procedure, resulting in autoregressive spectrum estimator, can be applied.

9.3.2. Autoregressive spectrum

BROCKWELL, P.J., and DAVIS, R.A. (2006) point out that for any real-valued stationary process \{x_t\} with continuous spectral density \(f(\omega)\) it is possible to find both \(AR(p)\) and \(MA(q)\) processes whose spectral densities are arbitrarily close to \(f(\omega)\). For this reason, in some sense, \{x_t\} can be approximated by either \(AR(p)\) or \(MA(q)\) process. This fact is a basis of one of the methods of achieving a consistent estimator of the spectrum, which is called an autoregressive spectrum estimation. It is based on the approximation of the stochastic process \{x_t\} by an autoregressive process of sufficiently high order \(p\):

\[ x_t = \mu + (\phi_1 B + \cdots + \phi_p B^p)x_t + \epsilon_t \]  \[7.108\]

where \(\epsilon_t\) is a white-noise variable with mean zero and a constant variance. The autoregressive spectrum estimator for the series \(x_t\) is defined as:

\[ \hat{s}(\omega) = 10 \times \log^{10} \frac{\sigma^2_x}{2\pi[1 - \sum_{k=1}^{p} \phi_k e^{-ik\omega}]} \]  \[7.109\]

where:

\(\omega\) – frequency, \(0 \leq \omega \leq \pi\);

\(\sigma^2_x\) – the innovation variance of the sample residuals;

\(\phi_k\) – \(AR(k)\) coefficient estimates of the linear regression of \(x_t - \bar{x}\) on \(x_{t-k} - \bar{x}\), \(1 \leq k \leq p\).

The autoregressive spectrum estimator is used in the visual spectral analysis tool for detecting significant peaks in the spectrum. The criterion of visual significance, implemented in JDemetra+, is based on the range \(\hat{s}^{\text{max}} - \hat{s}^{\text{min}}\) of the \(\hat{s}(\omega)\) values, where \(\hat{s}^{\text{max}} = \max_k \hat{s}(\omega_k)\); \(\hat{s}^{\text{min}} = \min_k \hat{s}(\omega_k)\); and \(\hat{s}(\omega_k)\) is the \(k\)th value of autoregressive spectrum estimator.

The particular value is considered to be visually significant if, at a trading day or at a seasonal frequency \(\omega_k\) (other than the seasonal frequency \(\omega_{60} = \pi\)), \(\hat{s}(\omega_k)\) is above the median of the plotted values of \(\hat{s}(\omega_k)\) and is larger than both neighbouring values \(\hat{s}(\omega_{k-1})\) and \(\hat{s}(\omega_{k+1})\) by at least \(\frac{1}{2}\) times the range \(\hat{s}^{\text{max}} - \hat{s}^{\text{min}}\).

Following the suggestion of SOUKUP, R.J., and FINDLEY, D.F. (1999), JDemetra+ uses an autoregressive model spectral estimator of model order 30. This order yields high resolution of strong components, meaning peaks that are sharply defined in the plot of $\hat{s}(\omega)$ with 61 frequencies. The minimum number of observations needed to compute the spectrum is set to $n = 80$ for monthly data and to $n = 60$ for quarterly series while the maximum number of observations considered for the estimation is 121. Consequently, with these settings it is possible to identify up to 30 peaks in $\pi_k$ the plot of 61 frequencies. By choosing $\omega_k = \ldots 60$ for $k = 0, 1, \ldots, 60$ the density estimates are calculated at exact seasonal frequencies ($1, 2, 3, 4, 5$ and 6 cycles per year).

The model order can also be selected based on the AIC criterion (in practice it is much lower than 30). A lower order produces the smoother spectrum, but the contrast between the spectral amplitudes at the trading day frequencies and neighbouring frequencies is weaker, and therefore not as suitable for automatic detection.

SOUKUP, R.J., and FINDLEY, D.F. (1999) also explain that the periodogram can be used in the visual significance test as it has as good as those of the AR(30) spectrum abilities to detect trading day effect, but also has a greater false alarm rate.

9.4. Sliding spans

The sliding spans technique involves the comparison of the correlated seasonal adjustments of a given period obtained by applying the adjustment procedure to a sequence of two, three or four overlapping spans of data, all of which contain this period (month or quarter).

Each period which belongs to more than one span is examined to see if its seasonal adjustments vary more than a specified amount across the spans. For the multiplicative decomposition a seasonal factor is regarded to be unreliable if the following condition is fulfilled:

$$\max_{k \in N} S_t = \max_{k \in N} t(k) - \min_{k \in N} S_t(k) > 0.03$$

[7.110]

---

208 The false alarm rate is defined as the fraction of the 50 replicates for which a visually significant spectral peak occurred at one of the trading day frequencies being considered in the designated output spectra (SOUKUP, R.J., and FINDLEY, D.F. (1999)).


\[ \min_{k \in N} S_t(k) \]

where:

\( S_t(k) \) – the seasonal factor estimated from span \( k \) for month (quarter) \( t \); \( N \)

\( N_t \)

– \( \{ k: \text{month (quarter) } t \text{ is in the } k\text{-th span}\} \).

For the additive decomposition JDemetra+ uses the rule in equation 7.11 for checking for the reliability of the seasonal factor.
\[
S_t = \sqrt{\frac{\sum_{i} i^2}{n}}
\]

where:
\(n\) – number of observations of the original time series \(y_i\).

The month-to-month percentage change in the seasonally adjusted value from span \(k\) for month \(t\) is calculated as:

\[
A
MM_m(k) = \frac{A_m(k)_{m-1} - A(m(k)_{m-1})}{A}
\]

where:
\(A_m(k)\) – the seasonally (and trading day) adjusted value from span \(k\) for month \(t\);
\(MM_m(k)\) is considered unreliable if the statistics below is higher than 0.03.

\[
MM_{mmax} = \max_{k \in N1_t} MM_m(k) - \min_{k \in N1_t} MM_m(k) > 0.03
\]

where:
\(N1_t\) – \(k\): month \(t\) and \(t-1\) are in the \(k\)-th span).

The respective formula for the quarter-to-quarter percentage change in the seasonally adjusted value from span \(k\) for quarter \(t\) is calculated as:

\[
A
QQ_q(k) = \frac{AQ_q(k)_{q-1} - A(k_q-1)}{A}
\]

where:
\(A_q(k)\) – the seasonally (and trading day) adjusted value from span \(k\) for quarter \(q\).

\(QQ_q(k)\) is considered unreliable if the statistics below is higher than 0.03.

\[
QQ_{qmax} = \max_{k \in N1_q} QQ_q(k) - \min_{k \in N1_q} QQ_q(k) > 0.03,
\]

where:
\(N1_q\) – \(k\): quarter \(t\) and \(t-1\) are in the \(k\)-th span).

The respective diagnostic can be also performed for the trading days/working days component.
9.5. Revision history

Revisions are calculated as differences between the first (earliest) adjustment of an observation at time $t$, computed when this observation is the last observation of the time series (concurrent adjustment, denoted as $A_t|t$) and a later adjustment based on all future data available at the time of the diagnostic analysis (the most recent adjustment, denoted as $A_t|N$).

In the case of the multiplicative decomposition the revision history of the seasonal adjustment from time $N_0$ to $N_1$ is a sequence of $R_{tA|N}$ calculated in a following way:

$$ R_{tA|N} = 100 \times \frac{A_t|N}{A_t|t} - 1. $$  

[7.116]

The revision history of the trend is calculated in a similar way:

$$ R_{tT|N} = 100 \times \frac{T_t|N}{T_t|t} - 1. $$  

[7.117]

where $T_t|t$ is the concurrent estimation of trend at time $t$ and $T_t|N$ is the most recent estimation, performed at the time of the diagnostic analysis.

With an additive decomposition $R_{tA|N}$ is calculated in the same way if all values $A_t|t$ have the same sign. Otherwise differences are calculated as:

$$ R_{tA|N} = A_t|N - A_t|t. $$  

[7.118]

The analogous expression for the trend component is:

$$ R_{tT|N} = T_t|N - T_t|t. $$  

[7.119]

9.6. Tests

This item explains the tests performed in the time domain, used by JDemetra+ to assess the quality of the fit and to detect the seasonal fluctuations. It is divided into two parts. The first part focuses on the test performed on the residuals from the modelling performed by the means of the
RegARIMA or TRAMO models. The second part discusses seasonality tests used to identify seasonal fluctuations on different stages of the seasonal adjustment process.

### 9.6.1. Tests on residuals

The tests outlined here include: Doornik-Hansen test for normality of residuals distribution, the Durbin-Watson test, the Ljung-Box and Box-Pierce tests for autocorrelation of residuals, and the Wald-Wolfowitz tests for randomness of residuals.

#### 9.6.1.1. Doornik-Hansen test

The Doornik-Hansen test for multivariate normality (DOORNIK, J.A., and HANSEN, H. (2008)) is based on the skewness and kurtosis of multivariate data that is transformed to ensure independence. It is more powerful than the Shapiro-Wilk test for most tested multivariate distributions.\(^{211}\)

The skewness and kurtosis are defined, respectively, as: 
\[ s = \sqrt{\frac{m^3}{m^2}} \text{ and } k = \frac{m^4}{m^2} \]
where: 
\[ m_i = \frac{\sum_{i=1}^{n} x_i^i}{n} \]
and \( n \) is a number of (non-missing) residuals.

The Doornik-Hansen test statistic derives from SHENTON, L.R., and BOWMAN, K.O. (1977) and uses transformed versions of skewness and kurtosis.

The transformation for the skewness \( s \) into \( z_1 \) is as in D’AGOSTINO, R.B. (1970):

\[
\beta = \frac{3(n^2 + 27n - 70)(n + 1)(n + 3)}{(n - 2)(n + 5)(n + 7)(n + 9)}
\]

\[
\omega^2 = -1 + \sqrt{2(\beta - 1)}
\]

\[
\delta = \frac{1}{\sqrt{\log(\omega^2)}}
\]

\[
y = s \frac{\sqrt{(\omega^2 - 1)(n + 1)(n + 3)}}{12(n - 2)}
\]

\[
z_1 = \delta \log(y + \sqrt{y^2 - 1})
\]

The kurtosis $k$ is transformed from a gamma distribution to $\chi^2$, which is then transformed into standard normal $z_2$ using the Wilson-Hilferty cubed root transformation:

\[
\delta = (n - 3)(n + 1)(n^2 + 15n - 4)
\]

[7.125]

\[
(n - 2)(n + 5)(n + 7)(n^2 + 27n - 70)
\]

[7.126]

\[
a = \frac{6\delta}{(n - 7)(n + 5)(n + 7)(n^2 + 2n - 5)}
\]

[7.127]

\[
c = \frac{6\delta}{(n + 5)(n + 7)(n^3 + 37n^2 + 11n - 313)}
\]

[7.128]

\[
l = \frac{12\delta}{(n - 2)(n + 5)(n + 7)(n^2 + 2n - 5)}
\]

[7.129]

\[
\alpha = a + c \times s^2
\]

[7.129]

\[
\chi = 2l(k - 1 - s^2)
\]

[7.130]

\[
z = \sqrt{9\alpha \left( \frac{1}{9\alpha} - 1 + \frac{3}{2\alpha} \right)} \cdot z
\]

[7.131]

Finally, the Doornik-Hansen test statistic is defined as the sum of squared transformations of the skewness and kurtosis. Approximately, the test statistic follows a $\chi^2$ distribution, i.e.:

\[
DH = z_1^2 + z_2^2 \sim \chi^2(2)
\]

[7.132]

### 9.6.1.2. Durbin-Watson test

The Durbin-Watson statistic is defined by\(^\text{212}\):

\[
\sum_{t=2}^{N}(\hat{a}_t - \hat{a}_{t-1})^2
\]

[7.133]

\[
d = \frac{\sum_{t=1}^{N} \hat{a}^2 t}{\sum_{N=1}^{N} \hat{a}^2 t}
\]

where:

\(\hat{a}_t\) – residual from the model.

Since $\sum_{t=2}^{N_t} (\hat{a}_t - \hat{a}_{t-1})^2 \equiv 2 \sum_{t=1}^{N_t} \hat{a}_t^2 - 2 \sum_{t=2}^{N_t} \hat{a}_t \hat{a}_{t-1}$, then the approximation $d \equiv 2(1 - r_{z,1})$, where $r_{z,1} = \frac{\sum_{t=1}^{N} \hat{a}_t \hat{a}_{t-1}}{\sum_{t=1}^{N} \hat{a}_t^2}$ is the autocorrelation coefficient of the residuals at lag 1, is true.

The Durbin-Watson statistics is between 0 and 4. When the model provides an adequate description of the data, then $r_{z,1}$ should be close to 0 and therefore the Durbin-Watson statistics is close to 2. When the Durbin–Watson statistic is substantially less than 2, there is evidence of positive serial correlation, while when it is substantially greater than 2 it indicates that the successive error terms are, on average, much different in value from one another, i.e., negatively correlated.

More formally, to test for a positive autocorrelation at significance level $\alpha$, the Durbin-Watson statistics is compared to the lower ($d_{l,a}$) and upper ($d_{u,a}$) critical values:

- If $d < d_{l,a}$ there is statistical evidence that the error terms are positively autocorrelated.
- If $d > d_{u,a}$ there is no statistical evidence that the error terms are positively autocorrelated.
- If $d_{l,a} < d < d_{u,a}$ the test is inconclusive.

Positive serial correlation is serial correlation in which a positive error for one observation increases the chances of a positive error for another observation.

To test for negative autocorrelation at significance $\alpha$, the test statistic $(4 - d)$ is compared to the lower ($d_{l,a}$) and upper ($d_{u,a}$) critical values:

- If $(4 - d) < d_{l,a}$ there is statistical evidence that the error terms are negatively autocorrelated.
- If $(4 - d) > d_{u,a}$ there is no statistical evidence that the error terms are negatively autocorrelated.
- If $d_{l,a} < (4 - d) < d_{u,a}$ the test is inconclusive.

9.6.1.3. Ljung-Box test

The Ljung-Box Q-statistics are given by:

$$LB(k) = n \times (n + 2) \times \sum_{k=1}^{k} \frac{\rho_{a,k}^2}{n - k}$$

where:

- $\rho_{a,k}$ is the autocorrelation coefficient at lag $k$ of the residuals $\hat{a}_t$.
- $n$ is the number of terms in differenced series;
$K$ is the maximum lag being considered, set in JDemetra+ to 24 (monthly series) or 8 (quarterly series).

If the residuals are random (which is the case for residuals from a well specified model), they will be distributed as $\chi^2_{(2K-m)}$, where $m$ is the number of parameters in the model which has been fitted to the data.

The Ljung-Box and Box-Pierce tests sometimes fail to reject a poorly fitting model. Therefore, care should be taken not to accept a model on a basis of their results. For the description of autocorrelation concept see 7.9.

9.6.1.4. Box-Pierce test

The Box-Pierce Q-statistics are given by:

$$BP(k) = n \sum_{k=1}^{K} \rho_{a^2,k}$$

where:

$\rho_{a^2,k}$ is the autocorrelation coefficient at lag $k$ of the residuals $\tilde{a}_t$.

$n$ is the number of terms in differenced series;

$K$ is the maximum lag being considered, set in JDemetra+ to 24 (monthly series) or 8 (quarterly series).

If the residuals are random (which is the case for residuals from a well specified model), they will be distributed as $\chi^2_{(2K-m)}$ degrees of freedom, where $m$ is the number of parameters in the model which has been fitted to the data.

The Ljung-Box and Box-Pierce tests sometimes fail to reject a poorly fitting model. Therefore, care should be taken not to accept a model on a basis of their results. For the description of autocorrelation concept see 7.9.

9.6.1.5. Wald-Wolfowitz test (Runs test)

The Wald-Wolfowitz tests checks for randomness of the residuals. Generally, it examines the hypothesis that a series of numbers is random. For data centred around the mean the test calculates the number and length of runs. A run is defined as a set of sequential values that are either all above or below the mean. An up run is a sequence of numbers each of which is above the mean; a down run is a sequence of numbers each of which is below the mean.
The test checks if the numbers of up and down runs are distributed equally in time. Either too many runs or too few runs are unlikely in a real random sequence. The null hypothesis is that the values of the series have been independently drawn from the same distribution. The test also verifies the hypothesis that the length of runs is random.

9.6.2. Seasonality tests

This section presents the set of seasonality tests implemented in JDemetra+. For the tests that originate from the X-11 algorithm a detailed description and testing procedure is available in LADIRAY, D., and QUENNEVILLE, B. (1999).

9.6.2.1. Friedman test (stable seasonality test)

The Friedman test (also called a stable seasonality test) is a non-parametric method for testing that samples are drawn from the same population or from populations with equal medians. In the regression equation the significance of the month (or quarter) effect is tested. The Friedman test requires no distributional assumptions. It uses the rankings of the observations.

Seasonal adjustment procedures use the Friedman test for checking for the presence of seasonality. For the purpose of the test the time series \( x_t \) is presented as a matrix of data \( \{x_{ij}\}_{n\times k} \) with \( n \) rows (the blocks, i.e. number of years in the sample) and \( k \) columns (the treatments, i.e. either 12 months or 4 quarters, depending on the frequency of the data). This matrix is transformed into a new matrix \( \{r_{ij}\}_{n\times k} \) where the entry \( r_{ij} \) is the rank of \( x_{ij} \) within block \( i \). In other words, \( r_{ij} \) is the rank of the period \( j \) in the year \( i \).

The test statistic is given by:

\[
Q = \frac{SS_t}{SS_e}
\]  
[7.136]

where:

\[
SS_t = n \sum_{j=1}^{k} (r_{ij} - \bar{r})^2;
\]

\[
SS_e = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{k} (r_{ij} - \bar{r})^2 \frac{1}{n(k-1)} - r_{..}^2;
\]

\( r_{..} \) is the average ranks of each treatment (month or quarter) \( j \) within each block (year); \( \bar{r} \) \( = \frac{1}{2} (N + 1) \) is by construction the average rank.

The test statistic represents the variance of the average ranking across treatments \( j \) in relation to the total. Under the null hypothesis of no seasonality, all periods (months or quarters) can be treated equally.
For large $n$ or $k$, i.e. $n > 15$ or $k > 4$, the probability distribution of $Q$ can be approximated by that of a chi-squared distribution. Thus, the p-value is given by $P(\chi^2(k - 1) > Q)$. If the null hypothesis of no stable seasonality is rejected at the 1% significance level, then the series is considered to be seasonal and the outcome of the test is displayed in green.

This test uses the preliminary estimate of the unmodified Seasonal-Irregular component\(^{213}\) (for X13ARIMA-SEATS this time series is shown in Table B3). In this estimate, the number of observations is lower than in the final estimate of the unmodified Seasonal-Irregular component. Because of this, the number of degrees of freedom in the stable seasonality test is lower than the number of degrees of freedom in the test for the presence of seasonality assuming stability (see 4.4.3). For example, X-13ARIMA-SEATS uses a centred moving average of order 12 to calculate the preliminary estimation of trend-cycle. Consequently, the first six and last six points in the series are not computed at this stage of calculation. The preliminary estimation of the trend is then used for the calculation of the preliminary estimation of the unmodified Seasonal-Irregular.

9.6.2.2. Kruskal-Wallis test

The Kruskal-Wallis test is a non-parametric test used for testing whether samples originate from the same distribution. The null hypothesis states that all months (or quarters, respectively) have the same mean. The rejection of the null hypothesis of the Kruskal-Wallis test implies that at least one sample stochastically dominates at least one other sample. The test does not identify where this stochastic dominance occurs or for how many pairs of groups stochastic dominance occurs. Under the null hypothesis the test statistic follows a $\chi^2$ distribution. When this hypothesis is rejected, it is assumed that the values of a time series differ significantly between periods and the test results are displayed in green. The test is typically applied to $k$ groups of data $\{x_{ij}\}$ each one $j = 1, ..., k$ is composed of $n_j$ observations which are indexed by $i = 1, ..., n_j$. Each month (or quarter) groups all the observations available for a certain number of years. As opposed to the notation used in the Friedman test, the number of observations here is not necessarily equal for each group. The ranking of each data point, represented by variable $r_{ij}$, is defined differently from how it is in the Friedman test, since it considers all observables $N = n_1 + \ldots + n_g$, thereby ignoring group membership. The test statistic is given by:

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

The test statistic is slightly different than the Friedman statistic. The numerator is defined as $SS_t = (N - 1) \sum_{j=1}^{g} n_j (r_{ij} - \bar{r})^2$ and the denominator is equal to $SS_e$

$$= \frac{1}{n(k-1)} \sum_{j=1}^{g} \sum_{i=1}^{n_j} (r_{ij} - \bar{r})^2,$$

\(^{213}\) The unmodified Seasonal-Irregular component corresponds to the Seasonal-Irregular factors with the extreme values.
where: \( n_j \) is the number of observations in group \( j \) (corresponding to the periods); \( r_{\bar{j}} \) is the average of the absolute ranks of the data in group \( j \); \( r = \frac{1}{2} (N + 1) \) is by definition the average rank.

Under the null hypothesis that all groups are generated from the same distribution, the test statistic \( Q \) is approximated by a \( \chi^2 \) distribution. Thus, the p-value is given by \( P(\chi^2_{g-1} > Q) \). This approximation can be misleading if some of the groups are very small (i.e. fewer than five elements). If the statistic is not significant, then there is no evidence of stochastic dominance between the samples. However, if the test is significant then at least one sample stochastically dominates another sample.

9.6.2.3. Test for the presence of seasonality assuming stability
The test statistic and testing hypothesis are the same as for Friedman stable seasonality test. The test statistic is calculated for the final estimation of the unmodified Seasonal-Irregular Component (in the case of X-13ARIMA-SEATS this series is presented in Table D8).

9.6.2.4. Evolutive seasonality test (Moving seasonality test)
The evolutive seasonality test is based on a two-way analysis of variance model. The model uses the values from complete years only. Depending on the decomposition type for the Seasonal-Irregular component it uses [7.136] (in the case of a multiplicative model) or [7.137] (in the case of an additive model):

\[
|S_{Ij} - 1| = X_{ij} = b_i + m_j + e_{ij}
\]  

\[
|S_{Ij}| = X_{ij} = b_i + m_j + e_{ij}
\]

where:
- \( m_j \) – the monthly or quarterly effect for \( j \)-th period, \( j = (1, ..., k) \), where \( k = 12 \) for a monthly series and \( k = 4 \) for a quarterly series;
- \( b_j \) – the annual effect \( i \), \( (i = 1, ..., N) \) where \( N \) is the number of complete years;
- \( e_{ij} \) – the residual effect.

The test is based on the following decomposition:

\[
S^2 = S_A^2 + S_B^2 + S_R^2
\]

\[
S^2 = \sum_{j=1}^{k} \sum_{i=1}^{N} (\bar{X}_{ij} - \bar{X}_.)^2 - \text{the total sum of squares;}
\]

\[
S^2 = N \sum_{j=1}^{k} (\bar{X}_{..} - \bar{X}_j)^2 - \text{the inter-month (inter-quarter, respectively) sum of squares, which mainly measures the magnitude of the seasonality;}
\]
$S_B^2 = k \sum_{i=1}^{n} (X_i - \bar{X})^2$ – the inter-year sum of squares, which mainly measures the year-to-year movement of seasonality;

$S_R^2 = \sum_{i=1}^{N} \sum_{j=1}^{k} (\bar{X}_{ij} - \bar{X}_i - \bar{X}_j - \bar{X})^2$ – the residual sum of squares.

The null hypothesis $H_0$ is that $b_1 = b_2 = \ldots = b_N$ which means that there is no change in seasonality over the years. This hypothesis is verified by the following test statistic:

$$F_M = \frac{S_R^2}{(n-1)(k-1)}$$

which follows an $F$-distribution with $k - 1$ and $n - k$ degrees of freedom.

9.6.2.5. Test for presence of identifiable seasonality

This test combines the values of the $F$-statistic of the parametric test for stable seasonality and the values of the moving seasonality test, which was described above.

The test statistic is:

$$T = \left( \frac{F_S}{F_M} \right)^{\frac{3}{2}}$$

where $F_S$ is a stable seasonality test statistic and $F_M$ is moving seasonality test statistic. The test checks if the stable seasonality is not dominated by moving seasonality. In such a case the seasonality is regarded as identifiable. The detailed description of the test is available in LOTHIAN, J., and MORRY, M. (1978).

9.6.2.6. Combined seasonality test

This test combines the Kruskal-Wallis test along with test for the presence of seasonality assuming stability ($F_S$), and evaluative seasonality test for detecting the presence of identifiable seasonality ($F_M$). Those three tests are calculated using the final unmodified S-I component. The main purpose of the combined seasonality test is to check whether the seasonality of the series is identifiable. For example, the identification of the seasonal pattern is problematic if the process is dominated by highly moving seasonality$^{214}$. The testing procedure is shown in Figure 9.14.

---

9.6.2.7. Test on autocorrelation at seasonal lags

The test checks the correlation between the actual observation and the observations lagged by one and two years. In the case of a monthly time series the Ljung-Box Q-statistic for testing the autocorrelation at the first and second seasonal lags becomes:

\[
QS = n(n + 2) \left( \frac{\hat{\beta}_{12}^2}{n - 12} + \frac{\hat{\beta}_{24}^2}{n - 24} \right)
\]

Under \( H_0 \), which states that the data are independently distributed, the statistics follows a \( \chi^2(S) \) distribution, where \( S = 2 \). Thus, the p-values are given by \( P(\chi^2(S) > Q) \). As \( P(\chi^2(2)) > 0.05 = 5.99146 \) and \( P(\chi^2(2)) > 0.01 = 9.21034 \), \( QS > 5.99146 \) and \( QS > 9.21034 \) would suggest rejecting the null hypothesis at 95% and 99% significance levels, respectively.
In the case of a quarterly time series the Ljung-Box Q-statistic for the testing the autocorrelation at the first and second seasonal lags becomes:

\[
Q_S = n(n + 2)(\frac{\hat{\rho}^2}{n-4} + \frac{\hat{\mu}^2}{n-8})
\]  

[7.144]

A significant Q_S statistic should be used to conclude that there is seasonality only when the sign of the autocorrelation coefficients is consistent with such a hypothesis (i.e. for Q_S > 0). When Q_S < 0, autocorrelation is associated with a two-year cycle215.

9.6.2.8. F-test on seasonal dummies

The F-test on seasonal dummies checks for the presence of deterministic seasonality. The model used here uses seasonal dummies (mean effect and 11 seasonal dummies for monthly data, mean effect and 3 for quarterly data) to describe the (possibly transformed) time series behaviour. The test statistic checks if the seasonal dummies are jointly statistically not significant. When this hypothesis is rejected, it is assumed that the deterministic seasonality is present and the test results are displayed in green.

This test refers to Model-Based \(\chi^2\) and F-tests for Fixed Seasonal Effects proposed by LYTRAS, D.P., FELDPAUSCH, R.M., and BELL, W.R. (2007) that is based on the estimates of the regression dummy variables and the corresponding t-statistics of the RegARIMA model, in which the ARIMA part of the model has a form (0,1,1)(0,0,0). The consequences of a misspecification of a model are discussed in LYTRAS, D.P., FELDPAUSCH, R.M., and BELL, W.R. (2007).

For a monthly time series the RegARIMA model structure is as follows:

\[
(1 - B)(y_t - \beta_1 M_{1,t} - \cdots - \beta_{11} M_{11,t} - \gamma X_t) = \mu + (1 - B)a_t
\]

[7.145]

where:

1 in month \(j = 1, \ldots, 11\)

\[M_{j,t} = \begin{cases} 
-1 & \text{in December} \\
0 & \text{otherwise} 
\end{cases}\]

\(y_t\) – the original time series;

\(B\) – a backshift operator;

\(X_t\) – other regression variables used in the model (e.g. outliers, calendar effects, user-defined regression variables, intervention variables);

\(\mu\) – a mean effect;

---

$a_t$ - a white-noise variable with mean zero and a constant variance.

In the case of a quarterly series the estimated model has a form:

$$(1 - B)(y_t - \beta_1 M_{1,t} - \cdots - \beta_3 M_{3,t} - \gamma X_t) = \mu + (1 - B)a_t$$

[7.146]

where:

1 in quarter $j = 1, \ldots, 3$

$M_{j,t} = \{-1$ in the fourth quarter - dummy variables; $0$ otherwise

One can use the individual $t$-statistics to assess whether seasonality for a given month is significant, or a chi-squared test statistic if the null hypothesis is that the parameters are collectively all zero. The chi-squared test statistic is $\chi^2 = \hat{\beta}' [Var(\hat{\beta})]^{-1} \hat{\beta}$ in this case compared to critical values from a $\chi^2(df)$-distribution, with degrees of freedom $df = 11$ (monthly series) or $df = 3$ (quarterly series). Since the $Var(\hat{\beta})$ computed using the estimated variance of $\alpha_t$ may be very different from the actual variance in small samples, this test is corrected using the proposed $F$ statistic:

$$F = \frac{\chi^2 \times n - d - k}{s - 1 \times n - d}$$

[7.147]

where $n$ is the sample size, $d$ is the degree of differencing, $s$ is time series frequency (12 for a monthly series, 4 for a quarterly series) and $k$ is the total number of regressors in the RegARIMA model (including the seasonal dummies $M_{j,t}$ and the intercept). This statistic follows a $F_{s - 1, n - d - k}$ distribution under the null hypothesis.

**9.6.2.9. Identification of seasonal peaks in a Tukey periodogram and in an autoregressive spectrum**

In order to decide whether a series has a seasonal component that is predictable (stable) enough, these tests use visual criteria and formal tests for the periodogram. The periodogram is calculated using complete years, so that the set of Fourier frequencies contains exactly all seasonal frequencies$^{216}$.

The tests rely on two basic principles:

- The peaks associated with seasonal frequencies should be larger than the median spectrum for all frequencies and;

---

$^{216}$ For definition of the periodogram and Fourier frequencies see 7.3.
The peaks should exceed the spectrum of the two adjacent values by more than a critical value.

In such a case, the test results are displayed in green. The statistical significance of each one of the seasonal peaks (i.e. frequencies $\frac{\pi}{6}, \frac{\pi}{3}, \frac{2\pi}{3}, \frac{5\pi}{6}$ corresponding to 1, 2, 3, 4 and 5 cycles per year) is also displayed.

Table 9.14: The seasonal and trading day frequencies by time series frequency.

<table>
<thead>
<tr>
<th>Number of months per full period</th>
<th>Seasonal frequency</th>
<th>Trading day frequency (radians)</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>$\frac{\pi}{6}, \frac{\pi}{3}, \frac{2\pi}{3}, \frac{5\pi}{6}$</td>
<td>$d, 2.714$</td>
</tr>
<tr>
<td>6</td>
<td>$\frac{\pi}{3}, \frac{2\pi}{3}$</td>
<td>$d$</td>
</tr>
<tr>
<td>4</td>
<td>$\frac{\pi}{2}, \frac{\pi}{7}$</td>
<td>$d, 1.292, 1.850, 2.128$</td>
</tr>
<tr>
<td>3</td>
<td>$\frac{\pi}{7}$</td>
<td>$d$</td>
</tr>
<tr>
<td>2</td>
<td>$\frac{\pi}{6}$</td>
<td>$d$</td>
</tr>
</tbody>
</table>

The calendar (trading day or working day) effects, related to the variation in the number of different days of the week per period, can induce periodic patterns in the data that can be similar to those resulting from pure seasonal effects. From the theoretical point of view, trading day variability is mainly due to the fact that the average number of days in the months or quarters is not equal to a multiple of 7 (the average number of days of a month in the year of 365.25 days is equal to $\frac{365.25}{12} = 30.4375$ days). This effect occurs $\frac{365.25}{12} \times \frac{1}{7} = 4.3482$ times per month: one time for each one of the four complete weeks of each month, and a residual of 0.3482 cycles per month, i.e. $0.3482 \times 2\pi = 2.1878$ radians. This turns out to be a fundamental frequency for the effects associated with monthly data. In JDemetra+ the fundamental frequency corresponding to 0.3482 cycles

\[ n \times \pi \times 42 \] per month is used in place of the closest frequency 60. Thus, the quantity 60 is replaced by $\omega_{42} = 0.3482 \times 2\pi = 2.1865$. The frequencies neighbouring $\omega_{42}$, i.e. $\omega_{41}$ and $\omega_{43}$ are set to, respectively, $2.1865 - \frac{1}{60}$ and $2.1865 + \frac{1}{60}$.

The default frequencies ($d$) for calendar effect are: 2.188 (monthly series) and 0.280 (quarterly series). They are computed as:

\[ \omega_{ce} = \frac{2\pi}{7} (n - 7 \times \lceil \frac{3}{7} \rceil), \]

where:
\( n = \frac{365.25}{s} \), \( s = 4 \) for quarterly series and \( s = 12 \) for monthly series.

Other frequencies that correspond to trading day frequencies are: 2.714 (monthly series) and 1.292, 1.850, 2.128 (quarterly series).

In particular, the calendar frequency in monthly data (marked in red in Figure 7.15) is very close to the seasonal frequency corresponding to 4 cycles per year \( \omega_{40} = \frac{2}{40} \pi = 0.20944 \).

![Figure 9.15: Periodogram with seasonal (grey) and calendar (red) frequencies highlighted.](image)

This implies that it may be hard to disentangle both effects using the frequency domain techniques.
9.7. The output items

The CSV, TXT and XLS outputs of JDemetra+ may contain the items shown in Table 9.15.

### Table 9.15: A list of output items of JDemetra+ CSV, TXT and XLS formats.

<table>
<thead>
<tr>
<th>Code</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>( y )</td>
<td>Original series</td>
</tr>
<tr>
<td>( y_f )</td>
<td>Forecasts of the original series</td>
</tr>
<tr>
<td>( y_{ef} )</td>
<td>Standard errors of the forecasts of the original series</td>
</tr>
<tr>
<td>( y_c )</td>
<td>Interpolated series</td>
</tr>
<tr>
<td>( y_{c,f} )</td>
<td>Forecasts of the interpolated series</td>
</tr>
<tr>
<td>( y_{c,ef} )</td>
<td>Standard errors of the forecasts of the interpolated series</td>
</tr>
<tr>
<td>( y_{lin} )</td>
<td>Linearised series (not transformed)</td>
</tr>
<tr>
<td>( l )</td>
<td>Linearised series (transformed)</td>
</tr>
<tr>
<td>( y_{cal} )</td>
<td>Series corrected for calendar effects</td>
</tr>
<tr>
<td>( y_{cal,f} )</td>
<td>Forecasts of the series corrected for calendar effects</td>
</tr>
<tr>
<td>( l_{f} )</td>
<td>Forecasts of the linearised series</td>
</tr>
<tr>
<td>( l_{b} )</td>
<td>Backcasts of the linearised series</td>
</tr>
<tr>
<td>( t )</td>
<td>Trend (including deterministic effects)</td>
</tr>
<tr>
<td>( t_{f} )</td>
<td>Forecasts of the trend</td>
</tr>
<tr>
<td>( sa )</td>
<td>Seasonally adjusted series (including deterministic effects)</td>
</tr>
<tr>
<td>( sa_{f} )</td>
<td>Forecasts of the seasonally adjusted series</td>
</tr>
<tr>
<td>( s )</td>
<td>Seasonal component (including deterministic effects)</td>
</tr>
<tr>
<td>( s_{f} )</td>
<td>Forecasts of the seasonal component</td>
</tr>
<tr>
<td>( i )</td>
<td>Irregular component (including deterministic effects)</td>
</tr>
<tr>
<td>( i_{f} )</td>
<td>Forecasts of the irregular component</td>
</tr>
<tr>
<td>( det )</td>
<td>All deterministic effects</td>
</tr>
<tr>
<td>( det_{f} )</td>
<td>Forecasts of the deterministic effects</td>
</tr>
<tr>
<td>( cal )</td>
<td>Calendar effects</td>
</tr>
<tr>
<td>( cal_{f} )</td>
<td>Forecasts of the calendar effects</td>
</tr>
<tr>
<td>( tde )</td>
<td>Trading day effect</td>
</tr>
<tr>
<td>( tde_{f} )</td>
<td>Forecasts of the trading day effect</td>
</tr>
<tr>
<td>( mhe )</td>
<td>Moving holidays effects</td>
</tr>
<tr>
<td>( mhe_{f} )</td>
<td>Forecasts of the moving holidays effects</td>
</tr>
<tr>
<td>( ee )</td>
<td>Easter effect</td>
</tr>
<tr>
<td>( ee_{f} )</td>
<td>Forecasts of the Easter effect</td>
</tr>
<tr>
<td>Code</td>
<td>Meaning</td>
</tr>
<tr>
<td>-----------</td>
<td>----------------------------------------------</td>
</tr>
<tr>
<td>omhe</td>
<td>Other moving holidays effects</td>
</tr>
<tr>
<td>omhe_f</td>
<td>Forecasts of the other moving holidays effects</td>
</tr>
<tr>
<td>out</td>
<td>All outliers effects</td>
</tr>
<tr>
<td>out_f</td>
<td>Forecasts of all outliers effects</td>
</tr>
<tr>
<td>out_i</td>
<td>Outliers effects related to irregular (AO, TC)</td>
</tr>
<tr>
<td>out_i_f</td>
<td>Forecasts of outliers effects related to irregular (TC)</td>
</tr>
<tr>
<td>reg</td>
<td>All other regression effects</td>
</tr>
<tr>
<td>reg_f</td>
<td>Forecasts of all other regression effects</td>
</tr>
<tr>
<td>reg_i</td>
<td>Regression effects related to irregular</td>
</tr>
<tr>
<td>reg_i_f</td>
<td>Forecasts of regression effects related to irregular</td>
</tr>
<tr>
<td>reg_t</td>
<td>Regression effects related to trend</td>
</tr>
<tr>
<td>reg_t_f</td>
<td>Forecasts of regression effects related to trend</td>
</tr>
<tr>
<td>reg_s</td>
<td>Regression effects related to seasonal</td>
</tr>
<tr>
<td>reg_s_f</td>
<td>Forecasts of regression effects related to seasonal</td>
</tr>
<tr>
<td>reg_sa</td>
<td>Regression effects related to seasonally adjusted series</td>
</tr>
<tr>
<td>reg_sa_f</td>
<td>Forecasts of regression effects related to seasonally adjusted series</td>
</tr>
<tr>
<td>reg_y</td>
<td>Separate regression effects</td>
</tr>
<tr>
<td>reg_y_f</td>
<td>Forecasts of separate regression effects</td>
</tr>
<tr>
<td>fullresiduals</td>
<td>Full residuals of the RegARIMA model</td>
</tr>
<tr>
<td>decomposition.y_lin</td>
<td>Linearised series used as input in the decomposition</td>
</tr>
<tr>
<td>decomposition.y_lin_f</td>
<td>Forecast of the linearised series used as input in the decomposition</td>
</tr>
<tr>
<td>decomposition.t_lin</td>
<td>Trend produced by the decomposition</td>
</tr>
<tr>
<td>decomposition.t_lin_f</td>
<td>Forecasts of the trend produced by the decomposition</td>
</tr>
<tr>
<td>decomposition.s_lin</td>
<td>Seasonal component produced by the decomposition</td>
</tr>
<tr>
<td>decomposition.s_lin_f</td>
<td>Forecasts of the Seasonal component produced by the decomposition</td>
</tr>
<tr>
<td>decomposition.i_lin</td>
<td>Irregular produced by the decomposition</td>
</tr>
<tr>
<td>decomposition.i_lin_f</td>
<td>Forecasts of the irregular produced by the decomposition</td>
</tr>
<tr>
<td>decomposition.sa_lin</td>
<td>Seasonally adjusted series produced by the decomposition</td>
</tr>
</tbody>
</table>
**decomposition.sa_lin_f**
Forecasts of the seasonally adjusted series produced by the decomposition

**decomposition.si_lin**
Seasonal-Irregular produced by the decomposition

**decomposition.x – tables.y**
For X-13ARIMA-SEATS only. Series from the X-11 decomposition (x = a, b, c, d, e; y=a1...)

**benchmarking.result**
Benchmarked seasonally adjusted series

**benchmarking.target**
Target for the benchmarking

---

The CSV matrix of JDemetra+ may contain:

<table>
<thead>
<tr>
<th>Code</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>span.start</td>
<td>Start of the series span</td>
</tr>
<tr>
<td>span.end</td>
<td>End of the series span</td>
</tr>
<tr>
<td>span.n</td>
<td>Length of the series span</td>
</tr>
<tr>
<td>espan.start</td>
<td>Start of the estimation span</td>
</tr>
<tr>
<td>espan.end</td>
<td>End of the estimation span</td>
</tr>
<tr>
<td>espan.n</td>
<td>Length of the estimation span</td>
</tr>
<tr>
<td>likelihood.neffectiveobs</td>
<td>Number of effective observations in the likelihood</td>
</tr>
<tr>
<td>likelihood.logvalue</td>
<td>Log likelihood</td>
</tr>
<tr>
<td>likelihood.adjustedlogvalue</td>
<td>Adjusted log likelihood</td>
</tr>
<tr>
<td>likelihood.ssqerr</td>
<td>Sum of the squared errors in the likelihood</td>
</tr>
<tr>
<td>likelihood.aic</td>
<td>AIC statistics</td>
</tr>
<tr>
<td>likelihood.aicc</td>
<td>Corrected AIC statistics</td>
</tr>
<tr>
<td>likelihood.bic</td>
<td>BIC statistics</td>
</tr>
<tr>
<td>likelihood.bicc</td>
<td>BIC corrected for length</td>
</tr>
<tr>
<td>residuals.ser</td>
<td>Standard error of the residuals (unbiased, TRAMOlike)</td>
</tr>
<tr>
<td>residuals.ser – ml</td>
<td>Standard error of the residuals (ML, X-13ARIMA-SEATS-like)</td>
</tr>
<tr>
<td>residuals.mean</td>
<td>Test on the mean of the residuals</td>
</tr>
<tr>
<td>residuals.skewness</td>
<td>Test on the skewness of the residuals</td>
</tr>
<tr>
<td>residuals.kurtosis</td>
<td>Test on the kurtosis of the residuals</td>
</tr>
<tr>
<td>residuals.dh</td>
<td>Test on the normality of the residuals (Doornik-Hansen tests)</td>
</tr>
<tr>
<td>residuals.lb</td>
<td>The Ljung-Box test on the residuals</td>
</tr>
<tr>
<td>Code</td>
<td>Meaning</td>
</tr>
<tr>
<td>------</td>
<td>---------</td>
</tr>
<tr>
<td>residuals.lb2</td>
<td>The Ljung-Box test on the squared residuals</td>
</tr>
<tr>
<td>residuals.seaslb</td>
<td>The Ljung-Box test on the residuals at seasonal lags</td>
</tr>
<tr>
<td>residuals.bp</td>
<td>The Box-Pierce test on the residuals</td>
</tr>
<tr>
<td>residuals.bp2</td>
<td>The Box-Pierce test on the squared residuals</td>
</tr>
<tr>
<td>residuals.seasbp</td>
<td>The Box-Pierce test on the residuals at seasonal lags</td>
</tr>
<tr>
<td>residuals.nrns</td>
<td>Test on the number of runs of the residuals</td>
</tr>
<tr>
<td>residuals.lrns</td>
<td>Test on the length of runs of the residuals</td>
</tr>
<tr>
<td>mstatistics.m1</td>
<td>The relative contribution of the irregular over three months span</td>
</tr>
<tr>
<td>mstatistics.m2</td>
<td>The relative contribution of the irregular component to the stationary portion of the variance</td>
</tr>
<tr>
<td>mstatistics.m3</td>
<td>The amount of period to period change in the irregular component as compared to the amount of period to period change in the trend-cycle</td>
</tr>
<tr>
<td>mstatistics.m4</td>
<td>The amount of autocorrelation in the irregular as described by the average duration of run</td>
</tr>
<tr>
<td>mstatistics.m5</td>
<td>The number of periods it takes the change in the trend-cycle to surpass the amount of change in the irregular</td>
</tr>
<tr>
<td>mstatistics.m6</td>
<td>The amount of year to year change in the irregular as compared to the amount of year to year change in the seasonal</td>
</tr>
<tr>
<td>mstatistics.m7</td>
<td>The amount of moving seasonality present relative to the amount of stable seasonality</td>
</tr>
<tr>
<td>mstatistics.m8</td>
<td>The size of the fluctuations in the seasonal component throughout the whole series</td>
</tr>
<tr>
<td>mstatistics.m9</td>
<td>The average linear movement in the seasonal component throughout the whole series</td>
</tr>
<tr>
<td>mstatistics.m10</td>
<td>The size of the fluctuations in the seasonal component in the recent years</td>
</tr>
<tr>
<td>mstatistics.m11</td>
<td>The average linear movement in the seasonal component in the recent years</td>
</tr>
<tr>
<td>mstatistics.q</td>
<td>Summary of the M-Statistics</td>
</tr>
<tr>
<td>mstatistics.q−m2</td>
<td>Summary of the M-Statistics without M2</td>
</tr>
<tr>
<td>diagnostics.quality</td>
<td>Summary of the diagnostics</td>
</tr>
<tr>
<td>diagnostics.basic checks. definition: 2</td>
<td>Definition test</td>
</tr>
<tr>
<td>diagnostics. basic checks. annual totals: 2</td>
<td>Annual totals test</td>
</tr>
<tr>
<td>diagnostics. visual spectral analysis. spectral seas peaks</td>
<td>Test of the presence of the visual seasonal peaks in SA and/or irregular</td>
</tr>
<tr>
<td>diagnostics. visual spectral analysis. spectral td peaks</td>
<td>Test of the presence of the visual trading day peaks in SA and/or irregular</td>
</tr>
<tr>
<td>diagnostics. regarima residuals. normality: 2</td>
<td>Test of the normality of the residuals</td>
</tr>
<tr>
<td>diagnostics. regarima residuals. independence: 2</td>
<td>Test of the independence of the residuals</td>
</tr>
<tr>
<td>diagnostics. regarima residuals. spectral td peaks: 2</td>
<td>Test of the presence of trading day peaks in the residuals</td>
</tr>
<tr>
<td>diagnostics. regarima residuals. spectral seas peaks: 2</td>
<td>Test of the presence of seasonal peaks in the residuals</td>
</tr>
<tr>
<td>diagnostics. residual seasonality. on sa: 2</td>
<td>Test of the presence of residual seasonality in the SA series</td>
</tr>
<tr>
<td>diagnostics. residual seasonality. on sa (last 3 years): 2</td>
<td>Test of the presence of residual seasonality in the SA series (last periods)</td>
</tr>
<tr>
<td>diagnostics. residual seasonality. on irregular: 2</td>
<td>Test of the presence of residual seasonality in the irregular series (last periods)</td>
</tr>
<tr>
<td>diagnostics. seats. seas variance: 2</td>
<td>Test on the variance of the seasonal component</td>
</tr>
<tr>
<td>diagnostics. seats. irregular variance: 2</td>
<td>Test on the variance of the irregular component</td>
</tr>
<tr>
<td>diagnostics. seats. seas/irr cross – correlation: 2</td>
<td>Test on the cross-correlation between the seasonal and the irregular component</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Code</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>log</td>
<td>Log transformation</td>
</tr>
<tr>
<td>adjust</td>
<td>Pre-adjustment of the series for leap year</td>
</tr>
<tr>
<td>arima. mean</td>
<td>Mean correction</td>
</tr>
<tr>
<td>arima. p</td>
<td>The regular autoregressive order of the ARIMA model</td>
</tr>
<tr>
<td>arima. d</td>
<td>The regular differencing order of the ARIMA model</td>
</tr>
<tr>
<td>arima. q</td>
<td>Regular moving average order of the ARIMA model</td>
</tr>
<tr>
<td>arima. bp</td>
<td>The seasonal autoregressive order of the ARIMA model</td>
</tr>
<tr>
<td>arima. bd</td>
<td>The seasonal differencing order of the ARIMA model</td>
</tr>
<tr>
<td>arima. bq</td>
<td>The seasonal moving average order of the ARIMA model</td>
</tr>
<tr>
<td>arima. phi(i)</td>
<td>Regular autoregressive parameter (lag=i, max i=3) of the ARIMA model</td>
</tr>
<tr>
<td>arima.th(i)</td>
<td>Regular moving average parameter (lag=i, max i=3) of the ARIMA model</td>
</tr>
<tr>
<td>arima.bphi(i)</td>
<td>Seasonal autoregressive parameter (lag=i, max i=1) of the ARIMA model</td>
</tr>
<tr>
<td>arima.bth(i)</td>
<td>Seasonal moving average parameter (lag=i max i=1) of the ARIMA model</td>
</tr>
<tr>
<td>regression.lp: 3</td>
<td>Coefficient and test on the leap year</td>
</tr>
<tr>
<td>regression.ntd</td>
<td>Number of trading day variables</td>
</tr>
<tr>
<td>regression.td(i): 3</td>
<td>Coefficient and test on the i&lt;sup&gt;th&lt;/sup&gt; trading day variable</td>
</tr>
<tr>
<td>regression.nmh</td>
<td>Number of moving holidays</td>
</tr>
<tr>
<td>regression.easter: 3</td>
<td>Coefficient and test on the Easter variable</td>
</tr>
<tr>
<td>regression.nout</td>
<td>Number of outliers</td>
</tr>
<tr>
<td>regression.out(i): 3</td>
<td>Coefficient and test on i&lt;sup&gt;th&lt;/sup&gt; the outlier (max i=16)</td>
</tr>
<tr>
<td>decomposition.seasonality</td>
<td>Presence of a seasonal component (1 – present, 0 – not present)</td>
</tr>
<tr>
<td>decomposition.trendfilter</td>
<td>The order of the trend filter</td>
</tr>
<tr>
<td>decomposition.seasfilter</td>
<td>The order of the seasonal filter</td>
</tr>
</tbody>
</table>

### 9.8. Benchmarking

Benchmarking<sup>217</sup> is a procedure widely used when for the same target variable the two or more sources of data with different frequency are available. Generally, the two sources of data do not agree, as an aggregate of higher-frequency measurements is not necessarily equal to the less-aggregated measurement. Moreover, the sources of data may have different reliability. Usually it is thought that less frequent data are more trustworthy as they are based on larger samples and compiled more precisely. The more reliable measurements are considered as benchmarks.

Benchmarking also occurs in the context of seasonal adjustment. Seasonal adjustment causes discrepancies between the annual totals of the seasonally unadjusted (raw) and the corresponding annual totals of the seasonally adjusted series. Therefore, seasonally adjusted series are benchmarked to the annual totals of the raw time series<sup>218</sup>. Therefore, in such a case benchmarking means the procedure that ensures the consistency over the year between adjusted and non-seasonally adjusted data. It should be noted that the 'ESS Guidelines on Seasonal Adjustment' (2015) do not recommend benchmarking as it introduces a bias in the seasonally adjusted data. Also the

---


U.S. Census Bureau points out that: *Forcing the seasonal adjustment totals to be the same as the original series annual totals can degrade the quality of the seasonal adjustment, especially when the seasonal pattern is undergoing change. It is not natural if trading day adjustment is performed because the aggregate trading day effect over a year is variable and moderately different from zero.* Nevertheless, some users may prefer the annual totals for the seasonally adjusted series to match the annual totals for the original, non-seasonally adjusted series. According to the ‘ESS Guidelines on Seasonal Adjustment’ (2015), the only benefit of this approach is that there is consistency over the year between adjusted and non-seasonally adjusted data; this can be of particular interest when low-frequency (e.g. annual) benchmarking figures officially exist (e.g. National Accounts, Balance of Payments, External Trade, etc.) where user needs for time consistency are stronger.

The benchmarking procedure in JDemetra+ is available for a single seasonally adjusted series and for an indirect seasonal adjustment of an aggregated series. In the first case, univariate benchmarking ensures consistency between the raw and seasonally adjusted series. In the second case, the multivariate benchmarking aims for consistency between the seasonally adjusted aggregate and its seasonally adjusted components.

Given a set of initial time series \( \{z_{i,t}\}_{i \in I} \), the aim of the benchmarking procedure is to find the corresponding \( \{x_{i,t}\}_{i \in I} \) that respect temporal aggregation constraints, represented by \( X_{i,T} = \sum_{t \in T} x_{i,t} \) and contemporaneous constraints given by \( q_{k,t} = \sum_{j \in J} w_{kj} x_{j,t} \), or, in matrix form: \( q_{k,t} = w_{k} x_{t} \).

The underlying benchmarking method implemented in JDemetra+ is an extension of Cholette’s method, which generalises, amongst others, the additive and the multiplicative Denton procedure as well as simple proportional benchmarking. Cholette’s method consists of minimising a quadratic penalty function, which, in general, can take different forms. Here the usual form is considered:

\[
\sum_{i,T} \left( \frac{x_{i,t} - z_{i,t} \rho}{\sigma^2} \right)^2 + \rho \left( \frac{x_{i,t-1} - z_{i,t-1}}{\sigma^2} \right)^2
\]

Notations used in the benchmarking procedure are:

- \( c \) – the number of periods of the initial time series included in the aggregated time series; it is assumed that the different time series start at the same date and all indices start at 0.
- \( t_c \) – the first multiple of \( c \) which is not greater than \( t \).

---

---
\[ x_t^c = \sum_{s=1}^{t_0} x_t \] and \[ x_t^c = \sum_{s=t_0}^t x_t \] – the cumulator variables, such as: \[ x_t^c = x_t^c + x_t \] and \[ X_T = \]

\[ X_{c \times (T+1) - 1}. \]

In the case of univariate benchmarking the quadratic penalty function of [7.149] corresponds, from a formal point of view, to the sum of the squared residuals generated by the autoregressive process \( \mu_t = \rho \mu_{t-1} + \epsilon_t \):

\[
(x_t | z_t) = |z_t \delta_{t+1} = \mu_t. \quad [7.150]
\]

To simplify the notation, \( |z_t|^2 \) will be denoted as \( \gamma_t \).

The benchmarked series can be derived from the state space model. In general, a state space model for a time series \( z_t \) consists of a measurement equation relating the observed data to a state vector \( \alpha_t \), and a transition equation that describes the evolution of the state vector over time.

In the context of the benchmarking procedure the state vector is:

\[
\alpha_t = (\delta_t \gamma_t), \quad [7.151]
\]

and the measurement equation, only defined for \( t = c \cdot T - 1 \), is:

\[
\delta_t^c = Z_t \cdot \alpha_t, \quad [7.152]
\]

where \( Z_t = (1 \quad \gamma_t) \).

The transition equation is defined as:

\[
\alpha_{t+1} = T_t \cdot \alpha_t + Q_t, \quad [7.153]
\]

where:

\[
T_t = \begin{cases} 
0 & 0 \\
0 & 0 & 1 
\end{cases} \quad \text{if } t + 1 = c \cdot T \\
\begin{cases} 
1 & \gamma_t 
\end{cases} \quad \text{otherwise}
\]

\[
Q_t = \begin{cases} 
0 & 0 \\
0 & 0 & 1 
\end{cases}
\]
The estimation method for this model implemented in JDemetra+ is the diffuse Kalman filter of Durbin-Koopman\textsuperscript{221}. In general, the Kalman filter is an algorithm for sequentially updating a linear projection for the system represented by a state space model.

The initial state vector is defined as:

\[
\alpha_{-1} = (0 \quad 1), \quad [7.154]
\]

and the initial (diffuse) variance matrices are:

\[
P_{\infty} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad P = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad [7.155]
\]

The aggregation constraints are modified as follows:

\[
X_{IT} = X_{IT} - \sum_{t \in T} z_{it} = \sum_{t \in T} \delta_{it}. \quad [7.156]
\]

Finally, the benchmarked series \((x_t)\) can easily be derived from the smoothed states \((\tilde{\alpha}_t)\):

\[
x_t = z_t + (0 \quad \gamma_t) \cdot \tilde{\alpha}_t. \quad [7.157]
\]

The state space representation of the multi-variate benchmarking model is obtained by juxtaposing the different matrices of the univariate models (one for each series involved in the model) and by adding, for each linear constraint, the corresponding measurement equation.

More precisely, the state vector is:

\[
\xi = \begin{pmatrix} \delta_- & \delta_t & \delta \\ \delta_t & \alpha & (\gamma_t \quad \delta_{nt} \quad \gamma_{nt}) \end{pmatrix}. \quad [7.158]
\]

After the constraints have been adapted to correspond to differences in comparison with the actual data, the vector of “observations” becomes:

\[
\tilde{q}^\text{t} = \begin{pmatrix} \vdots \quad 0 \quad \vdots \\ \vdots \end{pmatrix} \quad [7.159]
\]

\[
\tilde{q}_{kt}
\]

\textsuperscript{221} DURBIN, J., and KOOPMAN, S.J. (2001).
\[
\begin{align*}
\begin{bmatrix}
\gamma_{0t} & 0 & 0 \\
0 & \ddots & 0 \\
0 & 0 & 1
\end{bmatrix} & \text{if } t = c \cdot T - 1 \\
\end{align*}
\]

and the measurement matrix is:

\[
\begin{bmatrix}
1 & \gamma_{0t} & 0 & 0 \\
0 & \ddots & \vdots & 0 \\
0 & 0 & 1 & \gamma_{nt}
\end{bmatrix} \quad \text{for } t = c \cdot T - 1
\]

In other words, the vector \( \nu_t \) is composed of a sequence of contemporaneous constraints (for each \( t \) that does not correspond to the end of an aggregation period) and of temporal constraints\(^{222}\) (for each \( t \) that corresponds to the end of an aggregation period); the matrices of the measurement equation are defined accordingly. As mentioned above, the other matrices of the system are just the juxtaposition of the matrices defined in the univariate case.

By construction, the smoothed states contain MMSE estimates of the \( \delta_{it} \) that respect all the constraints of the model. The diffuse Kalman filter of Durbin-Koopman considers several features that can be used for an efficient implementation of the problem.

The JDemetra+ solution uses the following routines that are described in DURBIN, J., and KOOPMAN, S.J. (2001):

- The multivariate model is handled through its univariate transformation,
- The smoothed states are computed by means of the disturbance smoother.

\(^{222}\) Of course, the aggregation constraints must respect the contemporaneous constraints.
The performance of the resulting algorithm is highly dependent on the number of variables involved in the model ($\propto n^3$). The other components of the problem (number of constraints, frequency of the series, and length of the series) are much less important ($\propto n$).

From a theoretical point of view, it should be noted that this approach may handle any set of linear restrictions (equalities), endogenous (between variables) or exogenous (related to external values), provided that they don't contain incompatible equations. The restrictions can also be relaxed for any period by considering their "observation" as missing. However, in practice, it appears that several kinds of contemporaneous constraints yield unstable results. This is more especially true for constraints that contain differences (which is the case for non-binding constraints). The use of a special square root initialiser improves in a significant way the stability of the algorithm.

### 9.9. Autocorrelation function and partial autocorrelation function

#### Autocorrelation function

The correlation is a measure of the strength and the direction of a linear relationship between two variables. For time series the correlation can refer to the relation between its observations, e.g. between the current observation and the observation lagged by a given number of units. In this case all observations come from one variable, so similarity between a given time series and a $k$-lagged version of itself over successive time intervals is called an autocorrelation.

The autocorrelation coefficient at lag $k$ is defined as:

$$ (k) = \frac{\sum_{t=k+1}^{n} c}{\sum_{t=1}^{n} (x_t - \bar{x})^2} \quad \frac{\rho}{\Sigma \sigma^2} (x - \bar{x}) $$

[7.160]

where:

- $x_t$ – time series;
- $n$ – total number of observations; $\bar{x}$ – mean of the time series.

The set of autocorrelation coefficients $(k)$ arranged as a function of $k$ is the autocorrelation function (ACF). The graphical or numerical representation of the ACF is called autocorrelogram.
Figure 9.16: Autocorrelation function.

The autocorrelation function is a valuable tool for investigating properties of an empirical time series. The assessment of the order of an AR process simply from the sample ACF is not straightforward. While for a first-order process the theoretical ACF decreases exponentially and the sample function is expected to have the similar shape, for the higher-order processes the ACF maybe a mixture of damper exponential or sinusoidal functions which makes the order of the AR process difficult to identify. JDemetra+ displays the values of autocorrelation function for the residuals from the ARIMA model (see 4.2.7). The ACF graph (Figure 9.17), presents autocorrelation coefficients and the confidence intervals. If the autocorrelation coefficient is in the confidence interval, it is regarded as not statistically significant. Therefore, the user should focus on the values where the value of the ACF is outside the confidence interval. In JDemetra+ the confidence interval is indicated by two grey, horizontal, dotted lines.

Partial autocorrelation function

The partial autocorrelation is a tool for the identification and estimation of the ARIMA model. It is defined as the amount of correlation between two variables which is not explained by their mutual correlations with a given set of other variables.

Partial autocorrelation at lag $k$ is defined as the autocorrelation between $x_t$ and $x_{t-k}$ that is not accounted for by lags 1 through to $k-1$ which means that correlations with all the elements up to lag $k$ are removed. Following this definition, partial autocorrelation for lag 1 is equivalent to autocorrelation.

The partial autocorrelation function (PACF) is the set of partial autocorrelation coefficients $(k)$ arranged as a function of $k$. This function can be used to detect the presence of an autoregressive

---

process in time series and identify the order of this process. Theoretically, the number of significant lags determines the order of the autoregressive process.

![Partial autocorrelations](image)

Figure 9.17: Partial autocorrelation function.

The PACF graph (Figure 9.17: Partial autocorrelation function.), which is available from the Tools → Differencing menu presents partial autocorrelation coefficients and the confidence intervals. If the partial autocorrelation coefficient is in the confidence interval, it is regarded as statistically not significant. Therefore, the user should focus on the values where the absolute value of the PACF is outside the confidence interval. In JDemetra+ the confidence interval is indicated by two grey, horizontal, dotted lines.

9.10. Plugins

Installation of the new plugins from the local machine can be done from the Plugin functionality activated from the Tools menu.
To start the process, go to the Downloaded panel and click on the Add Plugins... option. Next the user should select the plugins from the folder in which the plugins have been saved and click OK button.

The new plugin is now visible in the panel.
Click on it and choose the **Install** button.

There is a wizard that allows the user to install the marked plugin(s). In the first step choose **Next** to continue or **Cancel** to terminate the process.
Next, mark the terms of agreements and choose **Install**.

Then the process is started.
After a while JDemetra+ informs about the outcome. Click **Finish** to close the window.

Once the process is finished, the newly installed plugin is integrated within the software. It is automatically activated and manifests its impact on the software. The picture below compares the view of the *Workspace* window before (on the left) and after (on the right) the installation of the NbDemetra-ODBC plugin.
The list of all installed plugins is displayed in the fourth panel. To modify the current settings mark the plugin (by clicking the checkbox in the Select column) and chose an action.

The following options are available:

- **Activate** – activates the marked plugin if it is currently inactive. The option is available for inactive plugins (see the picture below);
- **Deactivate** – deactivates the marked plugin if it is currently active. The option is available for active plugins (see the picture below);
- **Uninstall** – uninstalls the marked plugin.

Inactive plugins can be activated or uninstalled.
Active plugins can be deactivated or uninstalled.

There is a wizard that allows the user to activate/deactivate/uninstall the marked plugin(s). The example below illustrates the deactivation process. In the first step the user is expected to confirm or cancel the deactivation.

In the second step the user should decide if the software will be restarted immediately after the uninstallation is completed or not.
It is possible to delay the restart of the application, although the restart is necessary to complete the process.

Table 9.16: Default JDemetra+ plugins.

<table>
<thead>
<tr>
<th>Name</th>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>NbDemetra – Anomaly detection</td>
<td>SA core algorithms</td>
<td>Identification of outliers</td>
</tr>
<tr>
<td>NbDemetra – Spreadsheet</td>
<td>IO (Input/output)</td>
<td>Time series providers for spreadsheet (Excel, OpenOffice)</td>
</tr>
<tr>
<td>-------------------------</td>
<td>------------------</td>
<td>---------------------------------------------------------</td>
</tr>
<tr>
<td>NbDemetra – Common</td>
<td>IO (Input/output)</td>
<td>Common time series providers, like XML and TXT</td>
</tr>
<tr>
<td>NbDemetra – JDBC</td>
<td>IO (Input/output)</td>
<td>Time series provider for JDBC sources</td>
</tr>
<tr>
<td>NbDemetra – ODBC</td>
<td>IO (Input/output)</td>
<td>Time series provider for ODBC sources</td>
</tr>
<tr>
<td>NbDemetra – SDMX</td>
<td>IO (Input/output)</td>
<td>Time series provider for SDMX files</td>
</tr>
<tr>
<td>NbDemetra – Core</td>
<td>SA core algorithms</td>
<td>Encapsulation of the core algorithms</td>
</tr>
<tr>
<td>NbDemetra – UI</td>
<td>SA core algorithms</td>
<td>Basic graphical components</td>
</tr>
<tr>
<td>NbDemetra – Branding</td>
<td>SA core algorithms</td>
<td></td>
</tr>
<tr>
<td>NbDemetra – SA</td>
<td>SA core algorithms</td>
<td>Default SA framework, including TRAMO/SEATS and X-13ARIMA-SEATS. This implementation can lead to small differences in comparison with the original programs.</td>
</tr>
</tbody>
</table>

This list is displayed in the *Installed* panel, available from the *Plugin* functionality, activated from the *Tools* menu (Figure 9.18: Activation of the *Plugin* functionality from the *Tools* menu.).

![Figure 9.32: List of installed plugins.](image)
10. REFERENCES


Average Order’, *Biometrika*, 69, 81-94.

Cambridge University Press.


HARVEY, A.C., and KOOPMAN, S.J. (1992), ‘Diagnostic Checking of Unobseved-Components 

Adjustment’, *Journal of the American Statistical Association* 77, 63-70.

Adjusted Series to Annual Totals’, ASA proceedings.

samples’, *Biometrika* 76, 297–307.

University Press.


KAISER, R., and MARAVALL, A. (2000), ‘Notes on Time series Analysis, ARIMA Models and 

Seasonality and Current Trend-Cycle Assessment’, Proceedings of the Second International


