PRACTICAL GUIDE TO SEASONAL ADJUSTMENT WITH DEMETRA+

FROM SOURCE SERIES TO USER COMMUNICATION

UNITED NATIONS

ECE/CES/15

Note

PREFACE

This *Practical Guide* is the result of UNECE capacity-building activities in economic statistics for the countries of Eastern Europe, Caucasus and Central Asia. It suggests an overall process for performing seasonal adjustment and explains the related concepts. It brings together international recommendations for producing high quality time series, performing seasonal adjustment and disseminating the results. The *Guide* aims to assist statistical offices of Eastern Europe, Caucasus and Central Asia in producing economic statistics in a seasonally adjusted form, but it may also provide relevant insight into seasonal adjustment in general.

To assess the state of the economy and to make informed decisions on economic policy, correct and timely information has to be available about short-term economic development. However, since many economic phenomena such as production, income and employment are influenced by seasonal factors, simply relying on the raw, unadjusted statistical series may not give the right picture. Seasonally adjusted time series provide a clearer and more comparable measure of development which enables more timely detection of turning points. This is achieved by identifying and removing the seasonal pattern to reveal the underlying development.

Seasonal adjustment makes it easier to draw comparisons over time and to interpret the development in the series. It allows time series with different seasonal patterns to be compared between different industries or countries. It also makes the months or quarters of the year comparable with each other.

Demetra+ is seasonal adjustment software available free of charge on the Internet. It is maintained by Eurostat. Demetra+ includes two seasonal adjustment methods, X-12-ARIMA and TRAMO/SEATS. In the *Guide* we give instructions for using the current Demetra+ (version, 1.0.2+). For simplicity, the *Guide* focuses on using the TRAMO/SEATS method. However, this does not imply any preference between the two methods, both of which are commonly recommended. Further instructions for both methods are available from *the Demetra+ User Manual*. By taking a practical approach—especially for beginners—and covering basic issues from the quality of source series to user communication, this *Guide* complements training courses and user manuals.

The *Guide* draws on the international statistical recommendations and the work of Eurostat, ECB, OECD, UNSD and several national statistical offices and central banks on the methodology of short-term statistics, seasonal adjustment and data dissemination. It consolidates the main recommendations to construct a comprehensive overview of the guidance with relevance to the quality of seasonal adjustment. The referenced international recommendations include, in particular, *the ESS Guidelines on Seasonal Adjustment* (Eurostat, 2009), *the Demetra+ User Manual* (Grudkowska, 2011), *the International Recommendations for the Index of Industrial Production* (UNSD, 2010), *the Data and Metadata Reporting and Presentation Handbook* (OECD, 2007) and *Methodology of Short-term Business Statistics* (Eurostat, 2006).

In chapter 1 the *Guide* introduces the overall process of seasonal adjustment. This chapter may be useful for refreshing one's knowledge of seasonal adjustment or for simply learning the basics. Chapter 1 is a summary, and from there on, the chapters set out in detail the different steps of the process. Chapter 2 deals with assessing prerequisites; chapter 3 sets out the seasonal adjustment phase; chapter 4 presents several tools for quality assurance; and chapter 5 deals with issues relating to user communication. We recommend that beginners in seasonal adjustment start reading from chapter 2 onwards, and at the end, return to the summary provided by chapter 1.

Acknowledgements

This *Guide* builds on the materials prepared for training workshops on short-term economic statistics and seasonal adjustment for Eastern Europe, Caucasus and Central Asia. The World Bank co-financed the work that led to the development of the Guide.

The principal authors of the *Guide* are Anu Peltola (UNECE) and Necmettin Alpay Koçak (Turkish Statistical Institute). The work was initiated and supported by Carsten Boldsen (UNECE). The editing and formatting of the publication was carried out by Anu Peltola and Christina O'Shaughnessy (UNECE). Special thanks for valuable advice and contribution are extended to Augustín Maravall.

A UNECE training workshop for the countries of Eastern Europe, Caucasus and Central Asia in 2011 provided an opportunity to discuss practical problems on seasonal adjustment of official statistics in the region. Anu Peltola and Necmettin Alpay Koçak provided training on seasonal adjustment and time series methodology, while Petteri Baer (Statistics Finland) was in charge of training on dissemination and user communication. Thanks are due to all those who contributed to and participated in the training workshop.

The *Guide* draws heavily on the work by the members of the Eurostat-ECB high-level group of experts on seasonal adjustment who steered the development of Demetra+ software and the preparation of *the ESS Guidelines on Seasonal Adjustment*. *The Demetra+ User Manual* by Sylwia Grudkowska (National Bank of Poland) has provided essential information on the functionalities of Demetra+.

The expert group provided valuable comments on the draft. It comprises the following experts: Jean Palate (National Bank of Belgium), Ketty Attal-Toubert (INSEE), Robert Kirchner (Deutsche Bundesbank), Hans-Theo Speth (Destatis), Anna Ciammola (ISTAT), Sylwia Grudkowska (National Bank of Poland), Erika Földesi (Statistical Office of Hungary), Agustín Maravall (Central Bank of Spain), Itziar Alberdi Garriga (INE) and Gary Brown (ONS). International organizations are represented in the group by Andreas Hake and Mark Boxall of ECB; Daniel Defays, Pilar Rey del Castillo, Rainer Muthmann, Dario Buono, Gianluigi Mazzi, Jean-Marc Museux and Rosa Ruggeri-Cannata of Eurostat, as well as Frédéric Parrot (OECD) and Anu Peltola (UNECE). Ralf Becker and Julian Chow of the UNSD and Michael Richter of Deutsche Bundesbank also provided useful comments.

The *Guide* also commends the national and international experts who have worked for improving the quality of short-term statistics as well as their dissemination and revision practices.

List of abbreviations

AO	Additive outlier
ARIMA	Auto-Regressive Integrated Moving Average
BEA	Bureau of Economic Analysis (USA)
CSV	Comma Separated Values
CTRL	Control button
DESTATIS	Statistisches Bundesamt Deutschland
ECB	European Central Bank
ESS	European Statistical System
EU	European Union
EUROSTAT	Statistical Office of the European Union
IMF	International Monetary Fund
INE	Instituto Nacional de Estadística (National Statistical Institute of Spain)
INSEE	Institut National de la Statistique et des Etudes Economiques
ISTAT	Italian National Institute of Statistics
LS	Level shift
OECD	Organisation for Economic Cooperation and Development
ODBC	Open Database Connectivity
ONS	Office for National Statistics (of the United Kingdom)
SA	Seasonal adjustment
SCB	Statistics Central Bureau (of Sweden)
SDMX	Statistical Data and Metadata eXchange
SEASABS	Seasonal Analysis Australian Bureau of Statistics
SEATS	Signal Extraction in ARIMA Time Series
SNA	System of National Accounts
TC	Transitory change
TRAMO	Time series Regression with ARIMA noise, Missing observations and Outliers
TSW	TRAMO/SEATS for Windows
TS	TRAMO/SEATS
TXT	Text file
UNECE	United Nations Economic Commission for Europe
UNSD	United Nations Statistics Division
USCB	The X-12-ARIMA maintained by the U.S. Census Bureau
X-	Experimental (for example X-12)
XML	Extensible Markup Language
WK	Wiener-Kolmogorov test

Contents

Background

Introduction

The *Practical Guide to seasonal adjustment with Demetra+* responds to the need for support in the national statistical offices of Eastern European, Caucasus and Central Asian region in seasonal adjustment and time series methodology.

In 2008, the UNECE conducted a survey on the availability and international comparability of short-term statistics in the region. The lack of time series data and of seasonally adjusted series appeared as a particular problem in most countries. In part, international comparability of short-term economic statistics seemed to require some upgrading.

Only a few countries in the region had some experience in applying seasonal adjustment, and seasonally adjusted data were published rarely. According to the survey, the most commonly used seasonal adjustment method was X-12-ARIMA, and some used TRAMO/SEATS.

The most common statistics available in the seasonally adjusted form within the region were the gross domestic product, industrial production as well as exports and imports. Some countries also seasonally adjusted transport statistics, turnover in retail trade, employment and consumer prices.

All countries reported about their limited capacity in seasonal adjustment, but at the same time, they had plans to improve their capacity. Users had expressed interest in acquiring time series in a seasonally adjusted form. The lack of resources and training was the main obstacle for not doing so.

All national statistical offices in the region said that their organization needed assistance in seasonal adjustment. In particular, the countries mentioned the need for methodological material and training in Russian. They would also like to see an exchange of technical and methodological experience among statisticians in the region.

In September 2008, in Teheran, UNECE held a regional seminar on short-term statistics to bring together the national experts of the region to address key challenges related to short-term economic statistics and seasonal adjustment.

The seminar discussed constructing longer time series for industrial production based on the monthly or quarterly data releases by the Eastern European, Caucasus and Central Asian countries. The resulting series are now used as an example data set in this *Guide* and are presented in the form last confirmed by the countries in the start of 2011.

After the seminar in Teheran, the UNECE launched a capacity building programme on *New Challenges in Economic Statistics* for the period of 2010-2012. The co-financing provided by the World Bank has been essential for implementing the programme. The purpose was to promote regional cooperation in improving the quality of key economic statistics in the Eastern European, Caucasus and Central Asian countries.

The programme consists of training workshops and practical exercises. It discusses current problems and solutions in economic statistics. Some of the workshops address challenges with consumer price indices and some the implementation of the 2008 System of National Accounts (SNA). Two workshops focus on time series methodology and seasonal adjustment.

The seasonal adjustment workshop held in Astana, in March 2011, discussed the production and use of short-term statistics and the methods for improving their quality and timeliness for detecting turning points in the economy. It presented good practices in compiling time series, including how to treat changes in the business population and how to disseminate statistics taking into account pressing needs of users.

The workshop included practical exercises on seasonal adjustment with Demetra+ using the countries' own data. Many reported that they had started testing seasonal adjustment in order to improve international comparability and timeliness of their key economic indicators.

The UNECE asked the participants to continue testing seasonal adjustment with their own data and to present the results and problems in a follow-up workshop, in February 2012. This second seasonal adjustment workshop will deal with the problematic issues in seasonal adjustment based on the challenges faced by the participants.

As a result of the project, the UNECE offers at its website a set of training materials[1]. All training materials are available in English and in Russian. In addition, *the ESS Guidelines on Seasonal Adjustment* (Eurostat, 2009) were translated into Russian during the project.

UNECE continues to collect industrial production indices and carry out seasonal adjustment on behalf of the countries that do not yet release those data. The data are released monthly in the UNECE Statistical Database[2]. However, the UNECE encourages countries to take over seasonal adjustment of their national statistics.

Introducing seasonal adjustment and longer time series may require some rethinking of the statistical production. In some countries, time series have not been in the centre of attention of statistical production, as the focus has been on the current data. We hope that this Guide will assist the Eastern European, Caucasus and Central Asian countries in producing and releasing internationally comparable economic statistics in a seasonally adjusted form.

1 www.unece.org/stats

2 www.unece.org/stats/data

CHAPTER 1

Process of seasonal adjustment

Introduction

This chapter introduces the process of seasonal adjustment focusing on economic time series. We will take a look at the underlying statistical terminology and provide instructions for using Demetra+ software. For simplicity, we focus on one of the methods, TRAMO/SEATS. This chapter is a summary in which we aim to offer quick and concise instructions for seasonal adjustment.

Seasonal adjustment consists of four phases from the preparation of data to the publication of results (see table 1):

- Assessing prerequisites.
- Seasonal adjustment.
- Analysis of results.
- User communication.

Demetra+ software offers two different methods for seasonal adjustment, TRAMO/SEATS and X-12-ARIMA, which are the two most common methods. The National Bank of Belgium developed Demetra+ at the request of Eurostat. The Eurostat-European Central Bank (ECB) high-level group of experts on seasonal adjustment steered the development work. In 2009, the same group produced the European Statistical System *(ESS) Guidelines on Seasonal Adjustment*. The aim of creating the software was to offer a flexible tool that reflects the ESS Guidelines.

One of the main features of Demetra+ is to improve the comparability of these methods and provide common presentation tools for both of them. The software can handle either ad hoc analyses of one time series (single processing) or recurrent processes with multiple time series (multi-processing). Whilst the automatic adjustment by TRAMO/SEATS or X-12-ARIMA produces good results for most series, the quality diagnostics help confirm and refine the results.

Demetra+ software doesn't suggest a particular guided process for seasonal adjustment but offers several options for its users. Since it performs adjustments quickly and easily, without prior knowledge of seasonal adjustment theory, the user may end up confused about how it produced the results and how to interpret the quality diagnostics. Our practical tips, explanations and the suggested process may reduce the bewilderment and help unfold the user-friendly and flexible features of Demetra+.

Seasonal adjustment starts with checking the original data and preparing the data for adjustment. The quality of the raw data affects the quality of results to a large extent, e.g. accuracy, length of time series, quality of production methods and time consistency. Visual analysis tools are helpful in identifying outliers, missing values, volatility, presence of seasonality and breaks in the series.

In the second phase, the statistician makes decisions about how to treat unusual observations and calendar related effects and perform seasonal adjustment. In principle, seasonal adjustment includes either statistical modelling or smoothing of data with filters. The purpose is to separate repeating seasonal effects to reveal underlying development.

The varying number of holidays and working days within a month influences almost all economic time series. The user of statistics is not necessarily interested in knowing that production is higher because of two more working days in a given month. On the contrary, removing these kinds of calendar effects makes it easier to see the real increases or decreases in the level of activity. Using national calendars in seasonal adjustment, therefore, improves the results.

The third phase of the process analyses the results and the suitability of the identified statistical model in explaining the time series. The set of visual and numeric quality assessment tools of Demetra+ is useful for this purpose. Demetra+ provides summary diagnostics as a quick indication of the overall quality of adjustment and further details for refining the results.

Transparency about the methods and decisions increases the usefulness and comprehensibility of seasonally adjusted data. After thorough examination of the underlying raw data and the results, the decisions taken during the adjustment need to be documented for future use. Part of this internal documentation can feed into the user documentation. Lists of national holidays and events that have caused outliers are also useful for improving the estimation of the seasonal pattern.

Finally, seasonal adjustment aims at offering users of statistics a better service. One of the benefits of the seasonally adjusted data is the possibility to publish change from the previous month or quarter. Thus, seasonal adjustment provides a faster indication of changes in the level of economic activity.

Table 1 summarizes the process of seasonal adjustment and the steps of using Demetra+. The process starts with assessing the prerequisites for seasonal adjustment, i.e. by analysing the source time series and transferring the data into a suitable format for Demetra+. The second phase includes setting definitions, which TRAMO/SEATS or X-12-ARIMA apply during seasonal adjustment, and the third analyses the results with the tools offered by the software. Finally, the seasonally adjusted data are exported from Demetra+ and communicated to the users of statistics.

Table 1
The process of seasonal adjustment with Demetra+

Assess prerequisites	Open Demetra+
	Prepare source data
	Import data
	Check original series
Seasonal adjustment	Prepare calendars
	Select an approach
	Select regression variables
	Seasonally adjust
Analysis of results	Visual check
	Read quality diagnostics
	Refine and readjust
	Export data
User communication	Document choices
	Draft user documentation
	Prepare publication
	Support users

Assess prerequisites

Open Demetra+

To install Demetra+ software, go to:

circa.europa.eu/irc/dsis/eurosam/info/data/demetra.htm

Once installed, start the process with **Demetra+ by double clicking on Demetra.exe**. *The Demetra+ User Manual* and related user documentation contain detailed instructions for using the software.

The view to Demetra+ consists of different panels (image 1):

- *Browsers* panel (left) presents the time series.
- *Workspace* panel (right) shows information used or generated by the software.
- The central panel may contain several windows created by Demetra+. As it also displays the analyses, it is called *Results* panel in this *Guide*.
- *TS Properties* panel (bottom, left) displays the time series activated at the *Browsers* panel.
- *Logs* panel (bottom, right) contains log information describing activities performed by Demetra+.

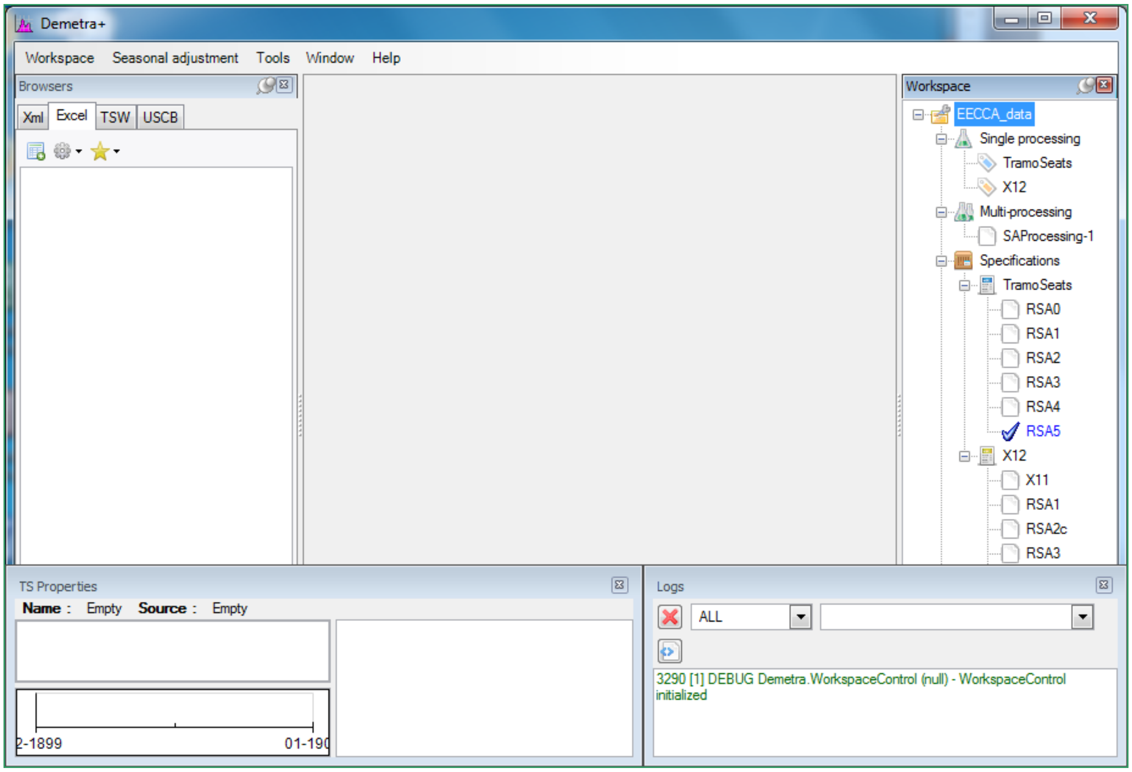
The user can move, resize and close panels as needed. Time series are easy to drag and drop between panels. You can re-open closed panels through the Main menu commands: **Workspace/View**. Demetra+ saves the chosen presentation mode for later use.

Prepare source data

Demetra+ provides an easy process for importing data from several types of files. It offers several simple solutions, such as the drag and drop facility or the clipboard. The various alternatives for dynamic data uploads include Excel, text and X-12-ARIMA software by the U.S. Census Bureau (USCB) files, Statistical Data and Metadata eXchange (SDMX), TRAMO-SEATS for Windows (TSW) and generic database drivers (ODBC) or WEB services and the XML format.

Deriving source data from Excel files is an easy solution which also offers a dynamic update possibility. This means that in the next seasonal adjustment of the same time series that includes new observations the software can read the same, updated source file. For more details about

Image 1
Panels of Demetra+



alternative ways to import source data, consult the *Demetra+ User Manual*.

To get started, prepare an Excel file with the data you wish to adjust. The file has to meet certain criteria to be a suitable input for Demetra+. The user can arrange the set of time series either vertically (image 2) or horizontally.

Format in a vertically structured Excel file should be as follows:

Image 2
A vertical data file

	A	B	C	D	E	F
1		AM	AZ	BY	BA	KZ
110	01-Jan-02	64.8	60.6	65.4	64.7	76.6
111	01-Feb-02	65.1	61.5	65.2	64.4	71.0
112	01-Mar-02	79.9	62.8	73.6	73.4	74.3
113	01-Apr-02	70.1	59.2	70.1	73.4	74.3
114	01-May-02	68.7	64.3	72.4	71.9	73.3
115	01-Jun-02	72.2	61.8	71.9	73.3	73.6
116	01-Jul-02	77.0	62.4	66.6	77.0	76.4
117	01-Aug-02	70.6	63.0	71.8	79.8	78.3
118	01-Sep-02	91.6	62.8	73.6	78.0	80.3
119	01-Oct-02	96.1	63.5	71.4	83.8	86.9
120	01-Nov-02	89.8	64.4	74.8	83.1	90.2
121	01-Dec-02	106.1	65.4	75.7	83.3	95.6
122	01-Jan-03	90.2	65.4	72.0	68.0	85.4
123	01-Feb-03	83.9	60.7	71.4	67.2	79.0

- True dates in the first column.
- Names of the series in the first row.
- Time series formatted as numbers.
- Empty top-left cell [A1].
- Empty cells (or -99 999) in the data zone correspond to missing values (except at the start and end of the series).

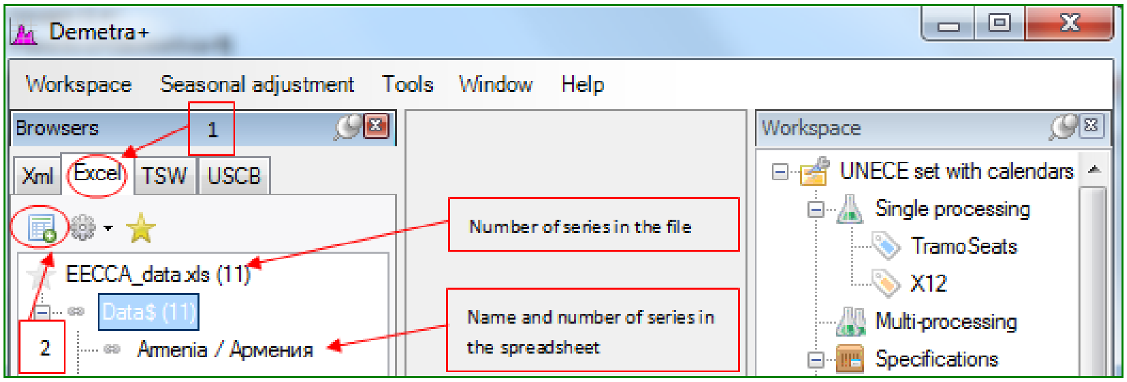
A horizontal presentation follows the same layout with the series names in the first column and periodicity in the first row.

Import data

Once the format of the Excel file corresponds to the previous instructions, the data can be imported. There are several alternative ways to import data.

The first option, set out in image 3, enables dynamic updates of the time series from the source file. In this case, the user doesn't have to import the data again when performing seasonal adjustment to the same data another time. Demetra+ will read the updated data from the

Image 3
Importing data to Demetra+ by reading an Excel file



original file as long as it remains in the same location with the same name.

Option 1
Read an Excel file

- **Click on the Excel tab** of the *Browsers* panel.
- **Click on the** **button** on the left in order to **Add** a source file.
- **Choose** the Excel file from your folders.

Option 2
Copy and paste data

- **Select** the entire data including the dates and titles in the Excel file. **Copy**.
- Click on the XML tab in the *Browsers* panel.
- Select **Paste** and the data will appear in the tree. If you have other files opened this way, the *Browsers* panel needs to be cleared by selecting **New**.
- You may **change the name** of the series in the tree if necessary by clicking on it twice (not by a double-click).
- **Save** the file in Demetra+.

If you copy and paste the data or drag and drop them into Demetra+ the automatic update is unavailable. However, for ad hoc seasonal adjustment this option is a quick choice.

You can **add, remove** or **clear** the contents of the *Browsers* panel **by right-clicking on a name** of a time series. By selecting Clear, you can remove the imported data from Demetra+.

Check original series

As the quality of the raw data affects the quality of seasonal adjustment, it is necessary to first check the original data: to consider the accuracy, length and consistency of time series and quality of production methods.

Visual analysis of time series is often helpful. It can, for instance, help identify the possible outliers, missing values, volatility, presence of seasonality and breaks in the seasonality or trend-cycle of the time series. Good documentation of the weaknesses of raw data helps share information with colleagues and users.

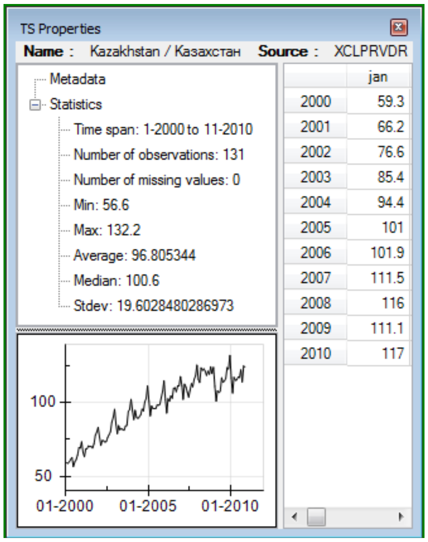
For seasonal adjustment, time series has to be at least 3 years long for monthly series and 4 years long for quarterly series. The quality of seasonal adjustment is likely to be higher with more than seven years of data.

On the other hand, very long series may not be ideal either, as they may not be consistent over time. The historical data may not reflect the seasonal pattern of the current data.

Seasonal adjustment requires discrete data for each period, month (or quarter), and it is not useful to adjust cumulative data.

In Demetra+, you will see the original time series and its basic properties in the *TS Properties* panel if you **click once on the name of the series** in the *Browsers* panel (image 4). The next image presents the time series, e.g. the number of observations, missing values and a graph of the original series.

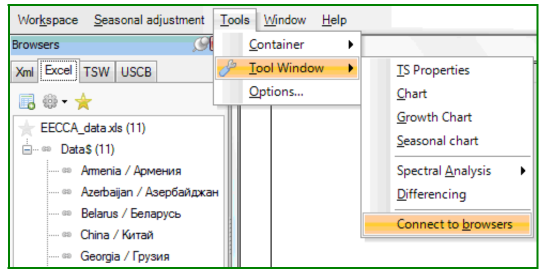
Image 4
Properties of the original time series



For a visual analysis, you can select **Tools/Container/Chart** or other options, such as a Growth chart or a Grid. Additional visual tools are available under **Tools/Tool window**. All these containers become active by dragging and dropping data from the *Browsers* (image 5) or from the Grid. From the Grid the data are selected by clicking on the top, left corner, cell A1

If you are working with a set of time series, first open a Tool from **Tools/Tool Window**, then select **Connect to Browsers** (image 6). This way the chart or grid will update every time you **click on a name** of a series in the *Browsers* panel. This provides a fast tool for visual checking of multiple time series one after the other. Note that you need to Connect to *Browsers* separately for each window, Chart or

Image 6
Visual checking of multiple series



Grid when they are active.

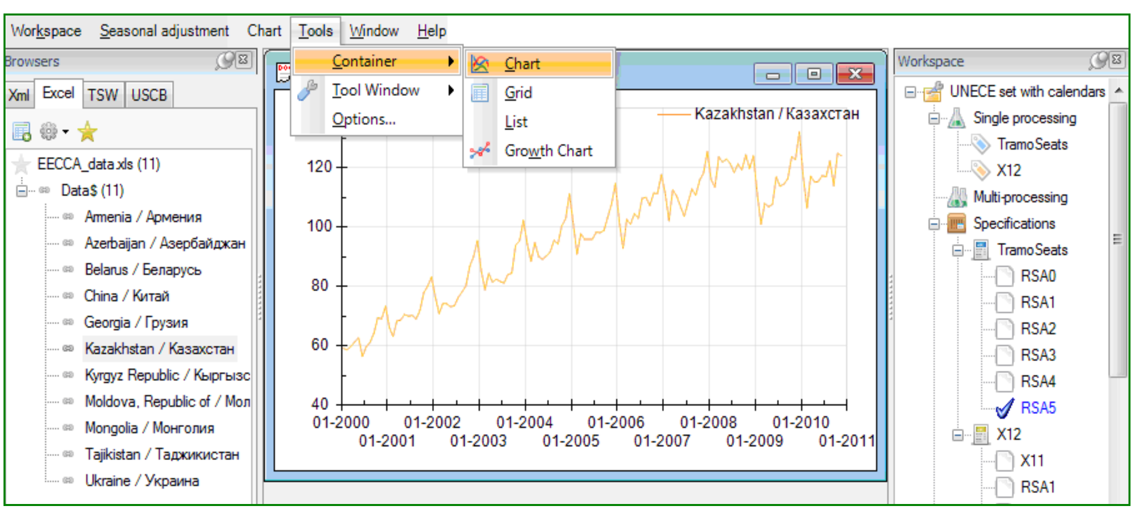
The Tool Window offers some tools specifically designed for analysing seasonality of a time series. With seasonal graphs, you can see quickly how different months or quarters differ from each other and how much the observations for each period differ between the years. To obtain a seasonal chart, select **Tools/Tool Window/ Seasonal Chart** and drag and drop the series to the window. In chapter 2 we'll provide further instructions for reading seasonal graphs.

Two important tools that deal with spectral analysis are the Periodogram and the Auto-Regressive Spectrum. They detect periodic components in a time series. Choose **Tools/Spectral Analysis/ Periodogram** to use these plots.

In a *Periodogram* and an *Auto-Regressive Spectrum*, seasonal frequencies are marked as grey vertical lines, while the purple lines correspond to trading day frequencies. Peaks at the seasonal or trading day frequencies indicate the presence of seasonality or trading day effects. Seasonality is a precondition for seasonal adjustment.

For further analysis, Demetra+ also provides a

Image 5
Using containers through drag and drop from the Browsers



Differencing tool for choosing the ARIMA model by determining the order of differencing. You can open it from **Tools/Tool Window/Differencing**.

A *stationary* time series is one whose statistical properties such as mean, variance and autocorrelation are constant over time. Most statistical forecasting methods assume that a time series can be transformed to make it approximately stationary. The purpose is to make the series easier to model. *Differencing* is a tool for making the time series stationary.

Seasonal adjustment

Prepare calendars

Using a list of national holidays is important for the quality of adjustment, as it improves the estimate of the calendar effects. This way, the calendar regression variables reflect country-specific situations more accurately. Before seasonal adjustment, make an effort to collect a time series of your national holidays. Demetra+ includes some regression variables for modelling predefined holidays but not for the national holidays of all countries. Thus, the user has to add them.

It is better to use official sources for the holidays when possible. To ensure good documentation, consider maintaining a separate list of holidays, outside of Demetra+, with explanations and exact dates.

Calendar effects, i.e. the effect of the varying number of holidays and working days influence most economic time series. The varying length of months, the number of different days appearing in a month, the composition of working and non-working days as well as different moving holidays may alter the level of activity described by a time series.

As an example of *moving holidays*, if the Easter holidays fall in March instead of April, the level of economic activity of these two months usually changes significantly. The moving Easter also influences quarterly series.

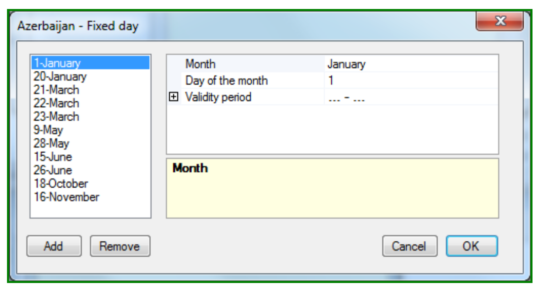
Calendar effects influence many time series, for example the retail sales index, but not all series. Quite often the pre-adjustment methods available in Demetra+, TRAMO and Reg-ARIMA, detect a difference in working days and non-working days. With some series, they may find a *trading day effect*, meaning that different days show different levels of activity. For instance, depending on the business branch, sales may be higher on Fridays than on Tuesdays.

To improve the seasonal modelling, TRAMO and Reg-ARIMA removes calendar effects before seasonal adjustment or the decomposition of the series. The pre-adjustment phase in Demetra+ is based on the functions of TRAMO and Reg-ARIMA methods.

To define holiday sets to Demetra+, **right-click on Calendars** and select **Edit** in the *Workspace* panel. A window appears that includes a tree for national calendars, composite calendars and chained calendars. To add a new calendar, click on the ... button and add.

Demetra+ provides alternative ways to define national holidays. First, you can select holidays from the pre-specified days, such as New Year, Christmas, Easter and May Day etc. However, these holidays may vary between countries, so be careful in choosing the correct dates. To start defining national holidays, **click on the row** in question. This opens a drop down list with options. **By clicking on the ... button** next to **Fixed days**, you can select the fixed holidays. **Add** the national holidays, and once finished, click **ok** (image 7).

Image 7
Adding national holiday calendars



Demetra+ offers a possibility for defining a validity period for each holiday. For example, if the date of a national holiday changes or the government decides to abolish a holiday, an option for limiting the duration of the holiday is useful. However, the validity periods should be used with caution. Demetra+ includes long-term corrections on the trading day variables when national calendars are used, but the correction doesn't take into account the validity period possibly leading to some seasonal effects in the variables. In more complex cases of moving holidays, you may also import the holiday regression variables to Demetra+ separately as we'll explain in chapter 3.

Select the approach

Demetra+ can process either one time series or multiple series at the same time. The user needs to select the seasonal adjustment approach, either X-12-ARIMA or TRAMO/SEATS.

TRAMO/SEATS applies seasonal adjustment filters based on statistical models to identify the components of time series, whereas X-12-ARIMA chooses from a priori designed moving average filters. The X-12-ARIMA and TRAMO/SEATS are the two most commonly used seasonal adjustment methods. The European Central Bank considers some level of combined use of the two methods as a preferable option (ECB, 2000).

In the *Workspace* panel, you can choose between the methods, X-12-ARIMA and TRAMO/SEATS. For practical instructions, this *Guide* refers to examples that use TRAMO/SEATS. For more instructions on X-12-ARIMA, you may consult the *Demetra+ User Manual*.

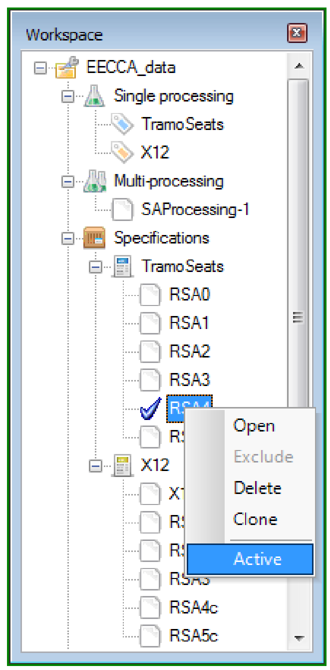
Define specifications and regression variables

Before any adjustments, select the specifications for TRAMO/SEATS or X-12-ARIMA to apply. You have five different readily programmed options in Demetra+. Alternatively, you can choose your own specifications by using the *Wizard*. In chapter 3 we'll explain how to use the *Wizard* for seasonal adjustment.

By definition, a *regression variable* is a variable that explains another variable, for instance an outlier or a calendar effect. By selecting regression variables, you take decisions about how to treat moving holidays, working days and trading days. TRAMO/SEATS and X-12-ARIMA remove the effect of the regression variables to estimate the seasonal pattern, but some of these effects are returned to the seasonally adjusted series, for example, outliers. You should apply a regression variable only if the series is influenced by the effect in question, and if the statistical tests prove it too.

You can start the analysis with the default specifications as shown in the *Workspace* panel (image 8). It is practical to choose first either the specification RSA4 or RSA5 for TRAMO/SEATS (table 2). For X-12-ARIMA one could start with RSA4(c) or RSA5(c). The difference between the RSA4 and RSA5 is that RSA4 performs a pre-test for the difference between working days and non-working days, while RSA5 looks for a difference between the days of the week. Clearly, the choice depends on the properties of the series. For instance, some series may not be influenced by trading day effects, for instance, quarterly data. Table 11, presented later in this *Guide,* includes also a description of the specifications for X-12-ARIMA. Choose the specification by **right-clicking on the option** and **select active** from the menu under TramoSeats.

Image 8
Specifications



Depending on your choice of the seasonal adjustment method, TRAMO or Reg-ARIMA tests

Table 2
Predefined specifications in TRAMO/SEATS

Name	Explanation
RSA0	level, Airline model
RSA1	log/level, outlier detection, airline model
RSA2	log/level, working days, Easter, outlier detection, airline model
RSA3	log/level, outlier detection, automatic model identification
RSA4	log/level, working days, Easter, outlier detection, automatic model identification
RSA5	log/level, trading days, Easter, outlier detection, automatic model identification

for trading and working day effect as well as leap year effect. It applies the regression variables only if the effects are significant. Demetra+ includes some pre-defined variables for modelling moving holidays (e.g. Easter), but not all moving holidays are included. If some regression variables for moving holidays are not pre-programmed, you can import the holiday regression variables to Demetra+ as we'll explain in chapter 3.

Seasonal adjustment is an exploratory process, where you learn by trying different sets of options. If the quality diagnostics give unsatisfying results after the first adjustment, another adjustment with different specifications could help. With the most important aggregate series, it is useful to try different sets of specifications in any case.

For other than default options, you may use the *Wizard* by selecting **Seasonal adjustment/Single analysis/Wizard.** Then it asks you to choose a series, a method and specifications. After the first seasonal adjustment with the predefined options, you can experiment with the *Wizard*. If need be, before launching the *Wizard*, you can add user-defined variables by means of the Main menu item **Workspace/Edit/User variables.**

Seasonally adjust

Seasonal adjustment separates the seasonal effects from a time series to reveal the underlying movement. To achieve this, it divides the series into its parts. The seasonally adjusted series and the seasonal component cannot be directly identified and extracted from a time series. As they have to be estimated, they are sometimes called the "unobserved components".

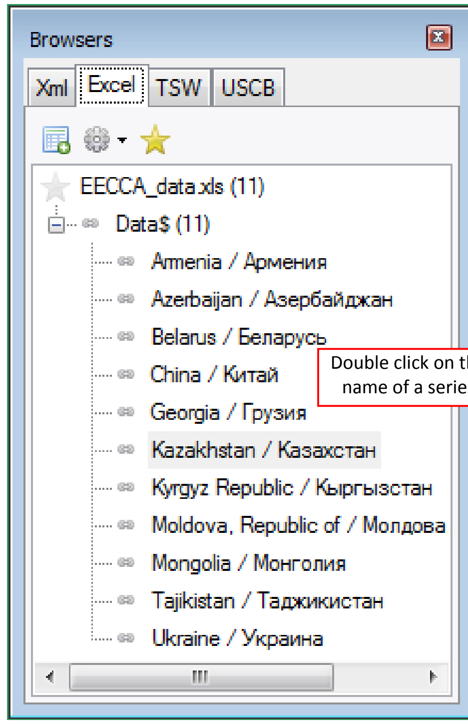
Seasonal adjustment estimates, identifies and separates the components of a time series. They include a seasonal, trend-cycle and irregular component. Sometimes a transitory component may also be identified. Once the components have been estimated, the irregular and transitory components can be put together into an irregular/transitory component. Depending on the series its components either sum up to form the original series or they are multiplied.

In chapter 2 we'll introduce the characteristics of the different components.

Single processing

There are several ways to launch seasonal adjustment of a single time series in Demetra+. Once you have selected the regression variables and the specification, you can launch the adjustment **by a double click on a series** in the *Browsers* panel (image 9). The results will appear in the middle panel. In chapter 3 we'll give further instructions e.g. for using the *Wizard.*

Image 9
Launching adjustment



Multi-processing

If you wish to adjust several series at once, you can create a multi-processing, through the Main menu **Seasonal adjustment/Multi-processing/New**. First, choose your **specification**, and second, **drag and drop the series** to the *Results* panel. Third, to start the adjustment, **select SAProcessing-n/Run** from the Main menu. In SAProcessing, "n" refers to the number of adjustments performed in Demetra+.

Any processing generated by double clicking is not saved in the *Workspace*. To save and later refresh it, select **SAProcessing-n/Add to Workspace** at the Main menu. Then save the *Workspace* by **Workspace/Save as** (image 10).

Image 10
Adding results to Workspace and saving them

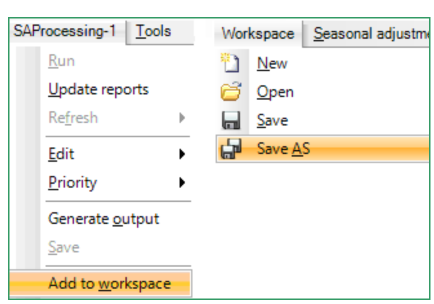


Image 11
Launching multi-processing in Demetra+

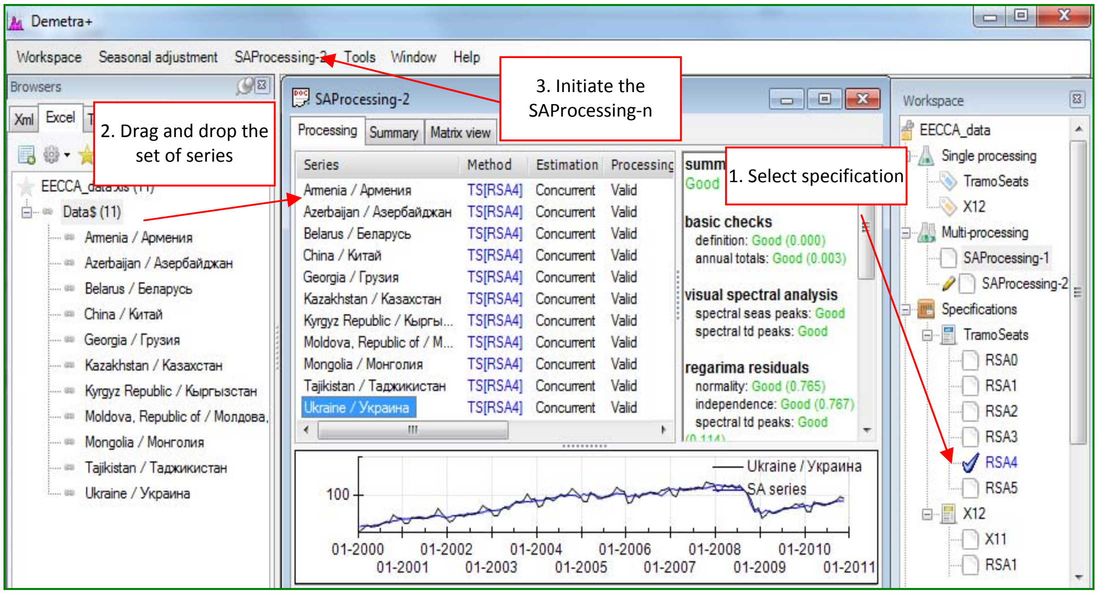


Image 11 summarizes the process for creating a multi-processing.

Analysis of the results

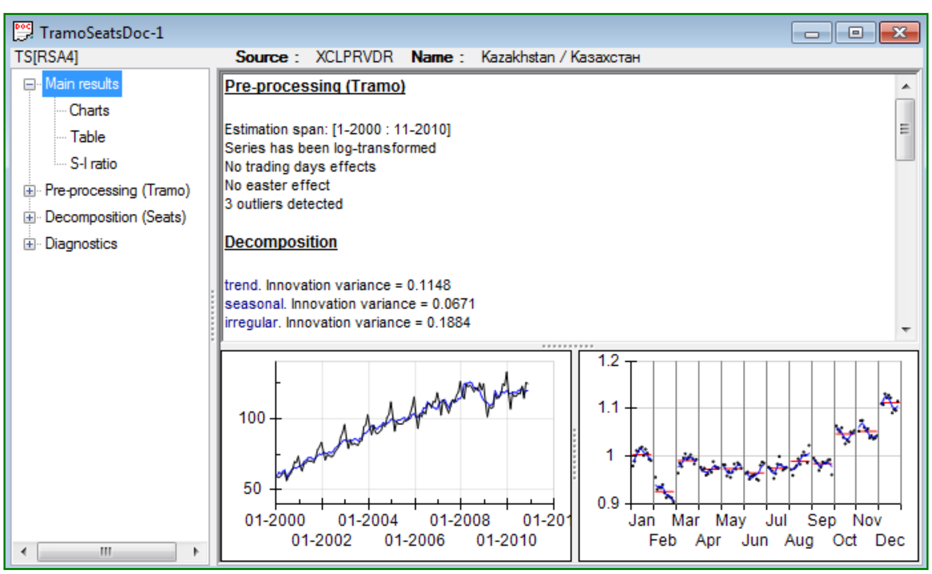
Visual check

After receiving the results of seasonal adjustment, you can make quality assessments by looking at the charts in Demetra+. The summary diagnostics under Main results give an indication of the overall quality of adjustment. If you wish to limit the amount of results displayed, you can do so by unselecting items for either TramoResults or X12Results at **Tools/Options/Default SA processing output**.

Demetra+ presents the results in the middle panel. In this example we'll present the results of TRAMO/SEATS. More information for X-12-ARIMA is available in the *Demetra+ User Manual*. The *Results* panel includes information divided into Main results, Pre-processing (TRAMO), Decomposition (SEATS) and Diagnostics. **By clicking on the Main results**, you can access the chart displaying the result series, the data table and a chart called the S-I ratio (image 12).

Demetra+ displays the basic results as charts

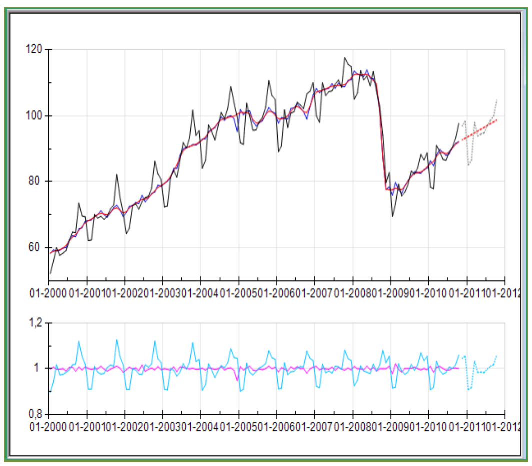
Image 12
Visualizing the seasonally adjusted results



including the original series (in black), seasonally adjusted series (in blue) and trend-cycle (in red) as well as the forecasts of all these series (image 13). Moreover, it provides a chart depicting the seasonal factor (in light blue in Demetra+) and the irregular component (in purple).

By looking at the lower chart (image 13), you can compare the **magnitude of seasonal variations** with the variations of the irregular component. If the irregular component is dominant, the seasonal component may be lost in the noise and TRAMO may not be able to identify a clear signal in the data. If so, you may say that the signal-to-noise ratio has become very small making the estimation problematic.

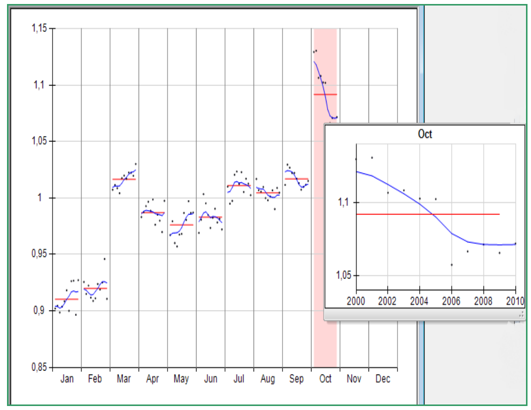
Image 13
Visualizing the components of time series



The **S-I ratio chart** is useful for analysing the development of the seasonal pattern, i.e. to detect unstable or **moving seasonal factors**. A sudden increase or decrease might imply a seasonal break, especially if it occurs for many consecutive months. A seasonal break means that the seasonal pattern of the series changes into a different one at a particular time. The methodology based on moving averages is sensitive to seasonal breaks, and they may complicate identifying trading day and Easter effects and fitting an ARIMA model.

By double clicking on a specific period, you can look at a certain month, i.e. October. The chart also helps detect months with higher variability (image 14). Seasonal breaks are problematic to treat. For instance, you could treat them with user-defined variables or adjust separately the two parts of the series. This would require several years of data before and after the break and is an option only for treating a historical break.

Image 14
S-I ratio depicts the seasonal pattern



Quality diagnostics

Models applied

Demetra+ offers the results of seasonal adjustment in the *Results* panel, including details about pre-processing and decomposition. The available quality diagnostics depend on the chosen seasonal adjustment method. The M-statistics of X-12-ARIMA are explained in the *Demetra+ User Manual*, and for further instructions on the interpretation of the SEATS results see Maravall and Pérez (2011).

The **Pre-processing** part shows the estimation span used, log-transformation, correction for trading days, the presence of the Easter effect and outliers. The corresponding heading on the left hand side in Demetra+ includes further details, e.g. the type of the applied ARIMA model, the regression variables and the dates and types of outliers. Demetra+ includes an analysis of the distribution of residuals and offers several tests on them.

Under the heading **Decomposition** of the *Results* panel, you'll find the applied decomposition model and more details e.g. about variance, autocorrelation and cross-correlation of the components. SEATS assumes that the components of a time series are orthogonal meaning that the theoretical components are uncorrelated. You can check this with the cross-correlations of the estimators and actual estimates of components.

Under the Main results for TRAMO/SEATS, you'll see the so-called innovation variance. The idea is to maximise the variance of the model for the irregular component to enable stable trend-cycle and seasonal components so that no additive white noise could be removed from them (image 15). This assumption is sometimes also called "canonical decomposition". By definition, the irregular component includes random fluctuations which cannot be attributed to the other components.

Image 15
Results panel

Pre-processing (Tramo)

Estimation span: [1-2000 : 11-2010]
Series has been log-transformed
No trading days effects
No easter effect
3 outliers detected

Decomposition

trend. Innovation variance = 0.1148
seasonal. Innovation variance = 0.0671
irregular. Innovation variance = 0.1884

Pre-processing – ARIMA model

TRAMO and Reg-ARIMA identify the most suitable ARIMA model and estimate the model parameters for each time series. Unless they find a specific ARIMA model in automatic model identification, they will apply the Airline model.

ARIMA models (p,d,q) are used for modelling and forecasting time series data. The ARIMA model includes three types of parameters: the autoregressive parameters (p), the number of differencing passes (d) and moving average parameters (q). A seasonal series usually has two sets of these parameter types: a regular component defined by (p,d,q) and a seasonal component (P,D,Q).

The Airline model (0,1,1)(0,1,1) is one of the most commonly used seasonal models. Box and Jenkins (1976) introduced it while studying a time series of the number of airline passengers.

The heading **Pre-processing** in the *Results* panel provides information about the statistical properties of the ARIMA model used in seasonal adjustment.

Pre-processing – Calendar effects

In Demetra+ regression variables can be defined for any frequency (e.g. monthly or quarterly). Demetra+ offers three options for treating calendar effects:

- **Trading Days** – Seven regression variables to model the differences in economic activity between all days of the week including the leap year effect.
- **Working Days** – Two regression variables to model the differences in economic activity between the working days (Monday to Friday) and non-working days (Saturday to Sunday) including the leap year effect.
- **None** – includes only one variable for the leap year effect.

By **double clicking on Pre-processing** in the *Results* panel, you'll see the calendar effects found, i.e. a trading day effect, an Easter effect or a leap year effect. From **Pre-processing/regressors** you can see further details about the regression variables applied (image 16).

Image 16
Results of pre-processing

ARIMA model [(0,1,1)(0,1,1)]

Parameter	*Value*	*Std error*	*T-Stat*	*P-value*
Th(1)	-0.1602	0.0944	-1.70	0.0925
BTh(1)	-0.6499	0.0907	-7.17	0.0000

Calendar effects
Trading days

Parameter	*Value*	*Std error*	*T-Stat*	*P-value*
Week days	0.0026	0.0004	5.79	0.0000
Sat/Sun (derived)	*-0.0064*	*0.0011*	*-5.79*	*0.0000*

Leap year

Leap year	*Value*	*Std error*	*T-Stat*	*P-value*
leap year	0.0308	0.0103	2.98	0.0035

Pre-processing – Outlier detection

Outliers may affect the reliability of the estimate for the seasonal pattern. There are at least three kinds of outliers. An additive outlier is a single point jump in the time series; a temporary or transitory change is a point jump followed by a smooth return to the original path; and a level shift is a more permanent change in the level of the series. A level shift is also sometimes referred to as a trend break. There may also be seasonal breaks which abruptly change the seasonal pattern.

TRAMO and Reg-ARIMA detect and replace outliers, i.e. abnormal values, before estimating the seasonal and calendar components. These include additive outliers (AO), transitory changes (TC) and level shifts (LS). Currently, Demetra+ does not include automatic options for identifying seasonal outliers and modelling for seasonal breaks. You can see the detected outliers by **double clicking on Pre-processing** in the *Results* panel (image 17).

Image 17
Results of pre-processing in outlier detection

Detected outliers

Parameter	Value	Std error	T-Stat	P-value
LS [10-2008]	-0.1555	0.0220	-7.08	0.0000
LS [11-2008]	-0.1201	0.0224	-5.36	0.0000
LS [9-2008]	-0.0781	0.0218	-3.58	0.0005
AO [12-2004]	-0.0537	0.0165	-3.26	0.0015

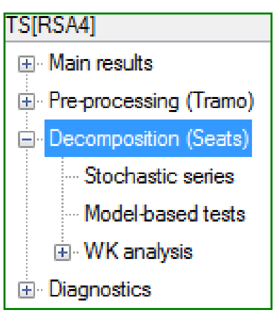
In Demetra+, TRAMO and Reg-ARIMA are responsible for performing statistical tests to identify outliers and set critical values for the tests by default. The critical value defines when to consider an observation an outlier. Although user can adjust the critical value and the outlier detection span, this kind of judgement requires experience. The ARIMA methods are sensitive to disturbances, breaks or outliers in the time series. To support the quality of seasonal adjustment, sensitive outlier detection is a safer choice. You can access the specifications by the Main menu and **Seasonal adjustment/Specifications**.

Decomposition Model

The decomposition performed by SEATS in Demetra+ assumes that all components in a time series, i.e. the seasonal, trend-cycle and irregular, are independent of each other. This applies also to the transitory component, which is sometimes identified and estimated. The method chooses a solution for identifying these components which maximises the noise of the model for the irregular component. The aim is that the trend-cycle and seasonal component are as smooth as possible.

Demetra+ provides the mathematical models of each component under **Decomposition** in the *Results* panel (image 18).

Image 18
Results of decomposition



You can see the result series, i.e. the seasonal, trend-cycle and irregular component, under **Stochastic series**. To test the validity of decomposition, Demetra+ offers some **Model-based tests**. There you can see if any cross-correlation exists between the components of the series. The theoretical components of a time series are assumed to be uncorrelated. Yet, the estimators of the components are usually correlated to some degree.

WK analysis (Wiener-Kolmogorov) includes advanced visual tools for analysing the decomposition further. In SEATS, the Wiener-Kolmogorov filters extract the components from the original series. The **phase effect** is one of the useful tools available for WK analysis. It reveals the possible time delay between the adjusted and the unadjusted series.

Diagnostics

To ensure the good quality of seasonal adjustment you should use the wide range of quality measures offered by Demetra+. The absence of residual seasonal and calendar effects and the stability of the seasonal component are among the most used tests. In this *Guide* we'll concentrate on the quality diagnostics of TRAMO/SEATS. The *Demetra+ User Manual* explains also the diagnostics for X-12-ARIMA. We'll offer more detail on using the diagnostic tools in chapter 4.

The quality diagnostics offered by Demetra+ include seasonality tests, spectral analysis, revision history, sliding spans and model stability. Under Diagnostics in the *Results* panel, you'll first see the **summary** diagnostics (image 19). They give a fast

Image 19
Main page of diagnostics

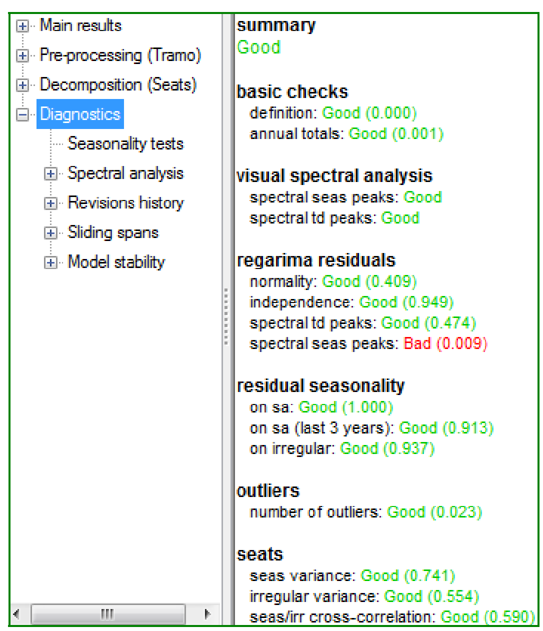


Table 3
Interpretation of the summary diagnostics from undefined to good

Value	***Meaning***
Undefined	The quality is undefined: unprocessed test, meaningless test, failure in the computation of the test
Error	There's an error in the results. The processing should be rejected (for instance, it contains aberrant values or some numerical constraints are not fulfilled
Severe	There's no logical error in the results but they shouldn't be accepted for some statistical reasons
Bad	The quality of the results is bad, following a specific criterion, but there's no actual error and you can use the results.
Uncertain	The result of the test is uncertain
Good	The result of the test is good

indication of the overall quality of the adjustment (table 3). Chapter 4 explains interpretation of these tools in more detail. The main page of **Diagnostics** summarizes the most relevant diagnostic results.

Diagnostics/Basic checks compare the **annual totals** of the original series and the seasonally adjusted series. The difference should be close to zero. The indicator called **definition** tests if the decomposition respects the mathematical relations of the different components and effects.

The **visual spectral analysis** of **Diagnostics** reveals remaining seasonality in a series where it shouldn't be present. The graphics for residuals, irregular component and seasonally adjusted series shouldn't show peaks at the seasonal or trading day frequencies (which appear as grey and purple vertical lines in Demetra+). Peaks at these lines would indicate the presence of seasonality and the need to find a better fitting model. **Diagnostics** shows the summary of these tests, and **Diagnostics/Spectral Analysis** presents the more detailed charts (image 20).

Diagnostics comprises information about **Reg-ARIMA residuals**, i.e. the part of data not explained by modelling. The analysis of Reg-ARIMA residuals constitutes an important test of the model. By definition, residuals shouldn't include any information, therefore, they should follow the normal distribution roughly, be random and independent.

Diagnostics presents summary results on the **residual seasonality** in order to reveal remaining seasonality in the seasonally adjusted series and the irregular component. For seasonality tests that are more detailed, go to **Diagnostics/Seasonality tests**. There you'll find the Friedman test, the Kruskall-Wallis test, the test for the presence of seasonality assuming stability, the evolutive seasonal test, the residual seasonality test and the combined seasonality test. Demetra+ will give

Image 20
Visual spectral analysis of the result series

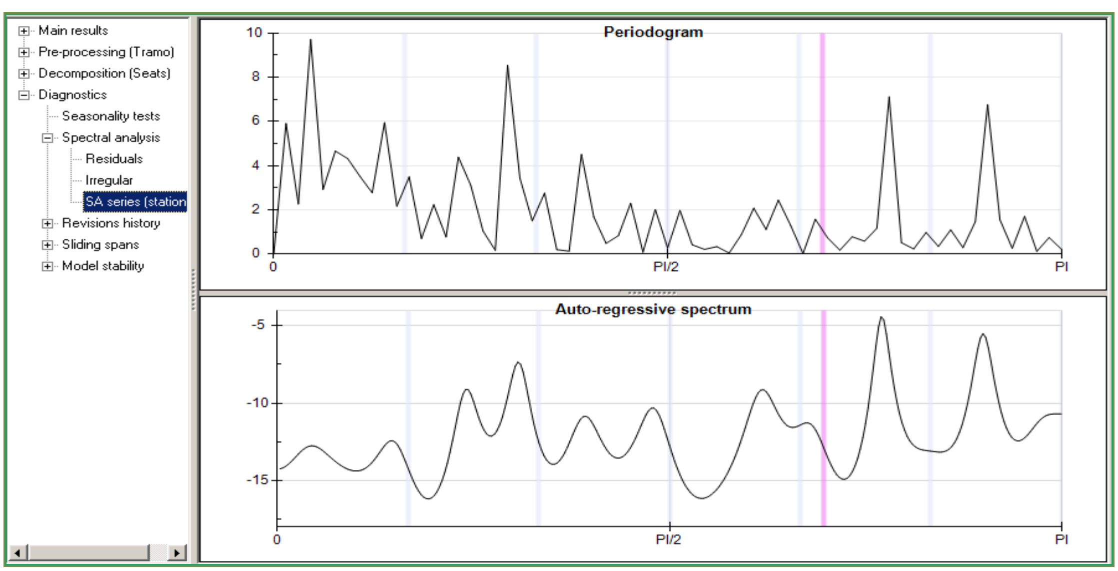
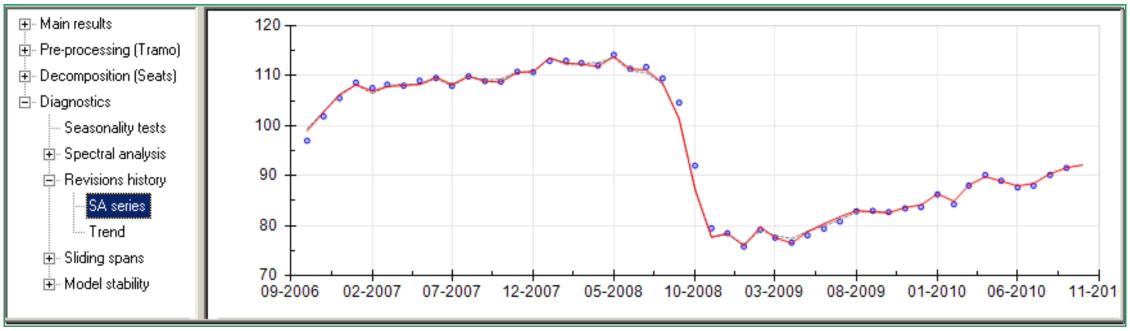


Image 21
Revision history as an indicator of the stability of adjustment



written conclusions on the test results, i.e. it states whether seasonality is present.

Diagnostics shows the number of **outliers** as an indicator of possibly weak stability of the process or a problem with the reliability of the data. If the number of outliers is high, it may compromise the quality of seasonal adjustment because the ARIMA model can't fit all observations into the model. However, with volatile series we have to accept a higher number of outliers.

Diagnostics presents summary statistics on the **seasonal variance** of the series. The cross-correlation table depicts the level of dependency between the components of the series and their estimators.

Diagnostics/Revision history includes useful charts for assessing the revisions of the seasonally adjusted and the trend-cycle series. The image 21 visualizes revisions to the seasonally adjusted series when new observations are added at the end of the series. The closer the initial observation dots are to the curve based on all available observations, the better the quality.

Diagnostics/Sliding spans analyses the stability of seasonal adjustment. It sets up time spans of 8 years, separated by 1 year, i.e. 2000 - 2008, 2001 - 2009 and 2002 - 2010. The SA series (changes) panel depicts period-to-period changes (image 22). If the decomposition of the series is additive, these sliding spans describe absolute differences, otherwise relative. You can consider values exceeding a three per cent threshold unstable.

Diagnostics/Model stability calculates the ARIMA parameters and coefficients of the regression variables (trading days, Easter etc.) for periods of eight years that slide one year at a time. The points on the chart in Demetra+ correspond to the different estimations. If the original time series is ten years long, there will be three dots on the chart, one for each period (2000-2008, 2001-2009, 2002-2010). The further the dots are from the line, the less stable the model.

Residuals

Residual analysis is one of the primary tools for verifying the appropriateness of the chosen seasonal model. Residuals are the portion of the data not explained by the model. The residuals shouldn't include any outstanding information or seasonality, i.e. they should be random.

Image 22
Sliding spans of the seasonal component

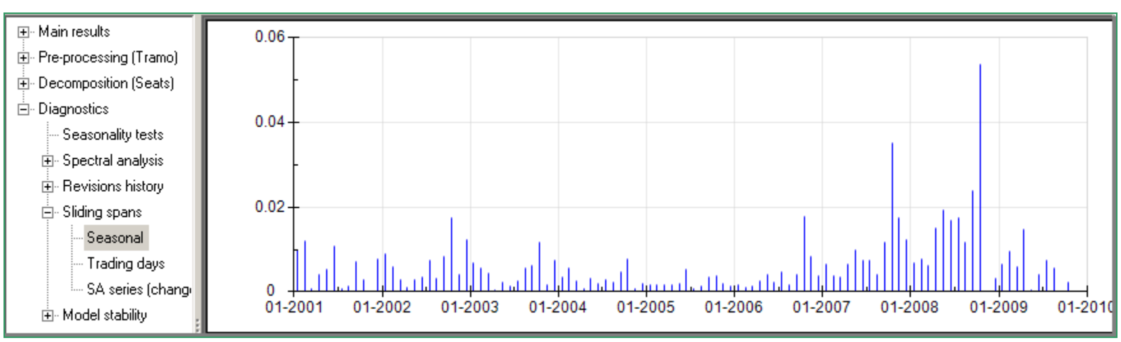
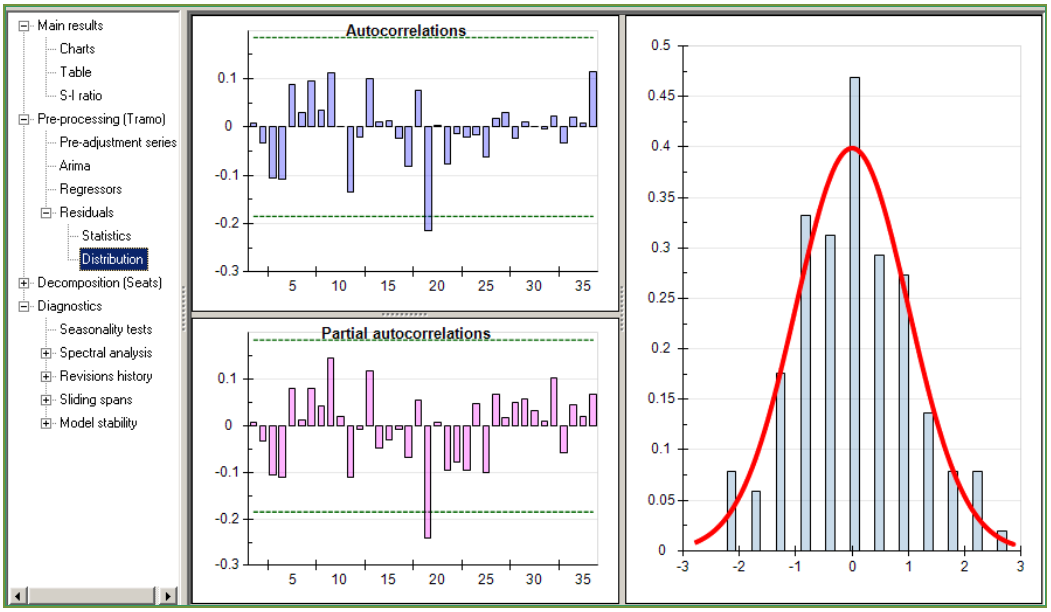


Image 23
Distribution of residuals



Autocorrelation refers to linear dependence between the values for different periods of a stationary variable. Tests for autocorrelation in the residuals are useful for detecting if a linear structure is left in the data. Residuals should not contain information, i.e. linear structures. The autocorrelation of residuals is useful in testing for a satisfactory fit of the ARIMA model to the data.

As the residuals shouldn't contain any information, they are supposed to roughly follow a normal distribution (image 23) and thus to have a mean of zero. You can see the distribution curves by selecting **Pre-processing/Residuals/Distribution**. In addition, Demetra+ provides statistics of residuals under **Pre-processing/Residuals/Statistics**.

For example, the Ljung-Box and Box-Pierce tests analyse the presence of remaining information or seasonality in the residuals. A green p-value refers to good results, yellow to uncertain and red to bad results. A p-value marked in red would indicate that Demetra+ has rejected the null hypothesis for this test. The result would be statistically significant meaning that one of the statistics on mean, normality, skewness or kurtosis would deviate from the normal distribution.

Mean

If the seasonal model fits the data, the residuals will follow normal distribution, and their mean should be zero. If not, TRAMO performs a mean correction to bring the mean of residuals to zero.

Kurtosis

A rejected null hypothesis signifies evidence of kurtosis in the residuals. Kurtosis is a statistical measure which describes the distribution of the observed data around the mean. It is a measure of how peaked or flat a distribution is relative to the normal distribution. The higher the value the more peaked the data.

Skewness

A rejected null hypothesis signifies some evidence of skewness in the residuals. This means that the residuals are asymmetrically distributed. It is a measure of how symmetrical a distribution is. A symmetrical distribution has a value of zero.

Normality

A rejected null hypothesis signifies asymmetry in the distribution of residuals and/or a pattern inconsistent with the normal distribution.

Ljung-Box on Residuals

A rejected null hypothesis signifies evidence of autocorrelation in the residuals. This indicates a remaining linear, unwanted structure in the series, i.e. outstanding information instead of only residual noise.

Box-Pierce on Residuals

This test examines evidence of autocorrelation in the residuals. A rejected null hypothesis signifies evidence of autocorrelation. As stated before, this indicates a remaining linear, unwanted structure in the series, i.e. outstanding information.

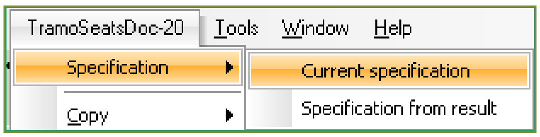
Refine and readjust

Refining the first results

Demetra+ offers many tools for refining the results. It's easy to change the specifications to see the effect on the quality of adjustment. In Demetra+, you can modify the regression variables or specifications and see the result immediately.

You can change the specifications by using the Main menu:

Image 24
Changing specification



Alternatively, you may, for instance, switch from the **specification** RSA5 to RSA4 **by dragging and dropping** the new specification from the *Workspace* to the middle panel and double clicking on the series to be adjusted (image 25).

This would mean applying one trading day variable for testing the difference between the working days and non-working days instead of applying six trading day variables to test for the difference between the days of the week. By double clicking on the series' name, you can adjust it again with the new specifications. The previous window and results will remain available.

In multiprocessing, you can edit any item of the processing **by double clicking its name** in the series list (image 26). The complete output, with the applied specification will open up. You can modify the specification in this new window, apply the new specification and save the results.

Refreshing the seasonally adjusted data

The need for seasonal adjustment usually repeats at regular intervals, e.g. monthly or quarterly.

When you re-open Demetra+, it will include the last *Workspace* in the **Main menu/Workspace**. If you have imported your data by the dynamic tools, Demetra+ will look for the updated series from the original location when you choose a refreshment strategy. The software offers many alternative refreshment strategies.

To adjust the series for the second time, go to the Main menu and select panel *Workspace*. If you were processing multiple time series, **double click** on the **SAProcessing-n** under the heading **Multi-processing**.

The previously adjusted data will appear in the middle panel. At the Main menu, you can start the second adjustment of these data by selecting **SAProcessing-n/Refresh** (image 27).

First, you need to select a refreshment strategy. You might apply current adjustment with fixed settings, or for instance, concurrent adjustment to re-estimate everything as during the first seasonal adjustment of a time series.

Image 25
Refining the specifications and readjusting

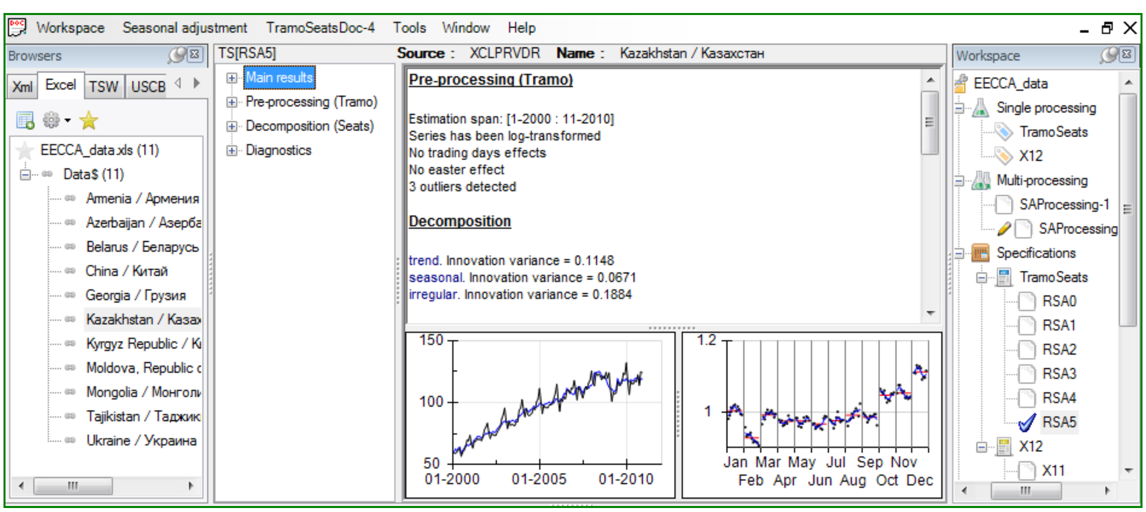
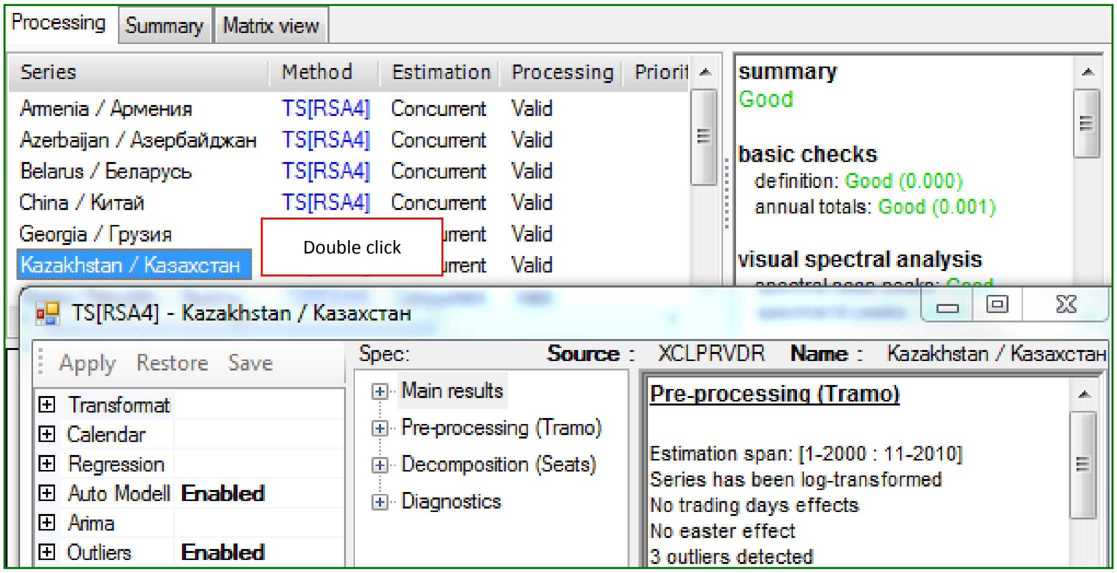


Image 26
Refining the specifications of an individual series in multiprocessing



Current adjustment means that the model, filters, outliers and calendar regression variables are re-identified and the respective parameters and factors re-estimated at review periods that have been set in advance. The method forecasts the seasonal model and its parameters and uses this information until the next review period which usually takes place once a year. Thus, current adjustment implies that the seasonal and calendar factors applied with new raw data in-between the review periods are fixed.

Partial concurrent adjustment usually means that the model, filters, outliers and calendar regression variables are re-identified once a year, but that the seasonal adjustment method re-estimates the respective parameters and factors every time new or revised observations become available. As described below, Demetra+ offers several modifications of the partial concurrent adjustment.

Concurrent adjustment means that the seasonal adjustment method re-identifies the model, filters, outliers and regression parameters with the respective parameters and factors every time new or revised data become available.

In Demetra+ you may select from three types of seasonal adjustment strategy:

- Current adjustment (partial): adjusts with fixed specification, user-defined regression variables can be updated.

- Partial concurrent adjustment:

  1. option: re-estimates coefficients, fixes the model, outliers and calendar effects.

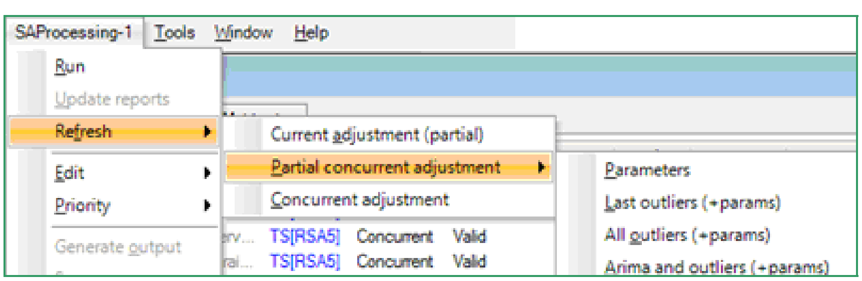
  2. option: same as the previous with re-estimation of the last outliers.

  3. option: same as the previous with re-estimation of outliers.

  4. option: same as the previous with re-estimation of the model.

- Concurrent adjustment: adjustment performed without any fixed specifications.

Image 27
Refreshing the seasonally adjusted data with new observations



In general, concurrent adjustment leads to more revisions, but at the same time, to more accurate results. Therefore, a compromise between concurrent and current adjustment is the most common seasonal adjustment strategy. The decision should take into account the properties of the series in question.

Stability is an important feature of the seasonally adjusted data. If TRAMO/SEATS or X-12-ARIMA suggests a different model in the annual update, you should examine the diagnostics to find out whether the model is notably better than the previous one. Assess also their effect on historical data and check the significance of the regression variables to identify any need for changes. In chapter 5 we'll address the issue of defining a seasonal adjustment strategy.

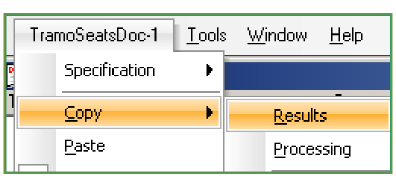
Export data

For further processing and publishing of data, you can export the results to other software. Demetra+ offers several alternative ways for exporting data and supports several kinds of outputs, e.g. Excel workbook or CSV files.

A simple alternative, e.g. for ad hoc seasonal adjustment, is to copy directly the results under **Main results/Table** in the *TS Properties* panel. By a **right-click on the corner** of the table, a window opens and you may select **Edit/Copy all**. Then you can paste the results to another programme, such as Excel, for further analysis.

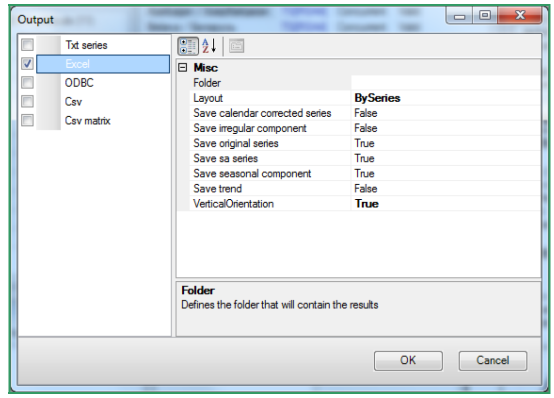
For single processing and smaller amounts of data, you may also directly copy the results by selecting **TramoSeatsDoc-n/Copy/Results** (image 28). Demetra+ gives a name to the folder in the Main menu depending on the method and the number of open adjustments. Next, you can paste the results to another file, e.g. to an empty Excel sheet.

Image 28
Copying results of the single processing



For exporting the results of multi-prosessing, go to the Main menu and select **SAProcessing-n/Generate output** (image 29). Demetra+ will save the Excel or the CSV file to the temporary folder if you don't specify a target folder.

Image 29
Exporting results of the multi-processing



User communication

Document choices

A sufficient amount of documentation helps ensure the quality of seasonal adjustment and provides the users with essential information. Up-to-date metadata should follow each release.

Systematic archiving of the resulting time series is important for later revision analysis. With the archived time series, you can improve the quality of seasonal adjustment in the longer term. You can analyse the behaviour of the seasonally adjusted series in the course of time, including during a turning point in the economy.

Transparency about methods and decisions includes offering enough metadata to enable the users to understand, and even to replicate, the seasonal adjustment. Users need metadata to assess the reliability of statistics and to use the seasonally adjusted data correctly.

The *ESS Guidelines on Seasonal Adjustment* include a metadata template. In addition, the *Data and Metadata Reporting and Presentation Handbook,* published by OECD, offers further details on data and metadata presentation.

Demetra+ provides a summary of quality diagnostics that you can store for the final seasonally adjusted series. For the most important series, you can copy the information of the summary statistics of *Results* panel, i.e. the first page of **Main results, Pre-processing, Decomposition and Diagnostics**. Documentation of the applied models, pre-processing choices and main diagnostics will leave you with the precise information that will be useful in the future, especially for re-estimating the seasonal model.

These documents are most useful, with a few words of conclusion regarding the quality of the resulting series.

To enable repetition of seasonal adjustment, you should provide the users of statistics with comprehensive information about the published figures. The user documentation should include details on the method and software used, decision rules, outlier detection and correction, events causing outliers, revision policy, description of trading day adjustment choices and contact information to the experts. Undoubtedly, there are different users with different needs. Therefore, you also need to prepare a non-technical and easily understandable explanation of seasonal adjustment.

To assess the quality of seasonal adjustment, you should select a set of quality indicators for publishing. You may want to prepare some metadata on quality issues, such as regarding the quality of the original data, the length of time series, average amount of revisions expected, the presence of strange values and a list of reasons for outliers.

Prepare publication

Ensure sufficient resources and enough time for analysing the results of seasonal adjustment before publishing the data for the first time. If you are introducing seasonal adjustment to a statistical news release for the first time, you may wish to redesign the content of data releases and the website.

In preparing the release, draft a document to explain the revisions of the seasonally adjusted data and prepare an advance schedule for data revisions. As much as the users of statistics prefer stability, they will value the accuracy of data as a result of revisions.

Revisions are an inevitable part of the seasonally adjusted data. The seasonally adjusted data revise for two main reasons, firstly due to the corrections and accumulation of raw data, and secondly, because of the revisions of the seasonal model caused by new observations. In addition, the two-sided filters and forecasts used to extend the series will revise when new observations accumulate.

Ideally a statistical release keeps the main message of the data simple and understandable. At the same time, it should deliver a comprehensive picture of the state of affairs. Therefore, the news release could include both some unadjusted data and seasonally adjusted data.

One of the benefits of the seasonally adjusted data is that it makes sense to publish change from the previous month. This provides a faster indication of changes in the economy if the underlying series is not too volatile. Additionally, users may be interested to know the change from the same month one year earlier: The working day adjusted or the original series is a good source for calculating this change per cent. Cumulative growth rates may be useful as additional information.

As the press releases should be simple, users may need further details from other sources. The website could offer some more disaggregated data, e.g. some regional or industry level data. Longer time series, e.g. the original series, the seasonally adjusted and the trend-cycle series could be available on the website.

You may wish to avoid presentation of the trend-cycle data in the press releases, as the end of the trend-cycle is unstable, and as it changes with new observations. However, the trend-cycle series may be good in visual presentations, for example, without the latest observations.

Support users

Before starting to publish seasonally adjusted data, you may benefit from discussions with the users of statistics. You can both learn from the users' needs, and at the same time, increase their knowledge and understanding of seasonal adjustment.

Before releasing seasonally adjusted data consider what kind of questions will arise among the users of statistics. Consider organising a seminar or training on key economic statistics and introduce your plans on seasonal adjustment as part of the training.

Be prepared to answer questions and be confident with the results but transparent with the possible quality issues. A set of documents for informing the users about seasonal adjustment and the related quality issues will make the process smoother. Furthermore, consider how to create easy access to the relevant metadata to all users.

In the end, seasonal adjustment aims at better service for users of statistics. Many offices performing seasonal adjustment are facing challenges in controlling the diversity of methods and decisions applied in seasonal adjustment. A clear seasonal adjustment policy would support establishing uniform practices.

CHAPTER 2

Assessing prerequisites

Introduction

This chapter addresses the first phase of the seasonal adjustment process: assessing the prerequisites; which includes preparing and checking the original data for seasonal adjustment. The chapter also explains the conditions a time series has to meet to be suitable for the adjustment.

To assess the prerequisites for seasonal adjustment, we have to study the:

- Quality of time series.
- Index calculation technique.
- Consistency of change measurement.
- Time series components.
- Effects influencing the series.

The quality of the raw data affects the quality of seasonal adjustment. You, therefore, need to first check the original data to consider the accuracy of data and the chosen production methods. The internationally recommended methods, definitions and classifications provide support in producing good quality statistics. To ensure accuracy of seasonal adjustment, one should allow revisions to the raw data and correct errors as part of the regular production process. Knowing your data and the factors that influence the series supports making decisions in seasonal adjustment.

The quality of seasonal adjustment benefits from the consistency of the raw time series, i.e. from the use of comparable methods, definitions and classifications as well as correction of breaks in the series. The methods of index calculation and change estimation affect the quality of the series. The inconsistencies, if any, should be identified and ideally removed before adjustment. In some countries, traditionally, time series have not been in the centre of attention of statistical production. Instead, the focus has been on the current data. In these cases, more attention needs to be paid on time series methodology, and therefore, this *Guide* also discusses some basic issues of index calculation.

The method of measuring change greatly influences the consistency of a series. Producing consistent time series starts from defining what the indicator should measure and how; secondly, you aim at correcting such changes in the population that don't describe a change in the measured variable. These refer to any changes not describing a change in the economic activity.

Producing consistent time series requires a good questionnaire design that allows to report revisions to previous data. A set of plausibility checks also helps identify outliers and correct errors effectively.

As seasonal adjustment aims to extract the seasonal effect from a time series, it's a precondition that seasonality is present in the raw data. TRAMO/SEATS or X-12-ARIMA divides the series into its main parts, namely the seasonal, trend-cycle and irregular component and sometimes also into the transitory component. The components fit together into the original series either by summing up or by multiplying. Once the components have been identified, seasonal adjustment removes the seasonal component from the raw data to compose the seasonally adjusted series.

Calendar related events often influence economic time series. For example, timing of holidays or seasonal habits may lead to different levels of economic activity in different months. If national holidays influence the series, the calendar effect can be more accurately defined if you use a national holiday calendar for estimating the calendar effect. Time series may also be influenced by interruptions, such as outliers, missing values, high volatility or seasonal breaks.

The Demetra+ software also has some prerequisites for the format of input data. Although flexible and allowing the use of several input formats and file types, it requires certain coding and structure from the input file. Demetra+ has a specific tool that is useful for repeating seasonal adjustment of the same source data. The software is able to read the revised and new figures from the original file if imported with the dynamic tools.

Visual analysis of time series is helpful for assessing the prerequisites for seasonal adjustment, and Demetra+ provides several useful tools for this. Visualizing the data may help confirm that the series fits the requirements of seasonal adjustment. Visual tools offer a quick view of the possible weaknesses of the data. The better you know the series before you embark on the seasonal adjustment, the easier it is to interpret the results afterwards.

Quality of time series

Seasonal adjustment is a method for transforming a time series into information that's easier to interpret. A time series is a collection of observations for a variable over time. It is measured at regular time intervals, e.g. monthly or quarterly. For seasonal adjustment, a time series has to be measured at time intervals shorter than one year because seasonal fluctuations are intra-annual and tend to repeat one year after the other.

Time series can be either measured as a stock or a flow series. Both types of series can be seasonally adjusted.

Stock series are measures of a variable taken at points in time. The monthly labour force survey is a usual example of a stock series. It takes stock of whether a person was employed in the reference period or not.

Flow series measure the activity over a given period. An example of a flow series could be the retail trade statistics: the daily sales are summed up to give a total value of sales for a month.

For data to be useful for time series analysis they should be comparable over time. That means that similar measurements should be taken over discrete and consecutive periods, i.e. every month or quarter. Seasonal adjustment requires discrete data for each period.

How we define and measure the variable should be consistent over time; and we should use the same definitions throughout the entire time series. If for any reason there are changes in definitions or measurements, we should adjust the time series to the changes so as to ensure consistency and to enable comparison over time and between countries.

Internationally recommended methods, definitions and classifications support the production of good quality time series. The *Methodology of Short-Term Business Statistics* (Eurostat, 2006), for example, provides useful guidance for many monthly and quarterly statistics. Index specific guides are also available, e.g. for the industrial production index, distributive trade statistics, producer prices and consumer prices.

The international guidelines for output indicators recommend revising the historical raw data to include newly accumulated observations and corrections. Since the quality of seasonal adjustment depends on raw data, you need to allow revisions to influence the underlying historical time series. Even one changed observation in the past could revise the seasonal pattern of a series. By correcting the seasonal model, the revised historical observations also improve the new estimates and, thus, the overall quality of seasonal adjustment.

For the adjustment, a monthly time series has to be at least three years long and include 36 observations. For quarterly data, the series should be four years long with 16 observations. As mentioned earlier, the quality of seasonal adjustment is likely to be better with more than seven years of data. The number of years of available data correspond to the number of examples for estimating the seasonal component.

Very long time series don't necessarily lead to higher quality seasonal adjustment. If the series aren't consistent in the way the data are measured or defined, it might be better to shorten them for the purpose of identifying the seasonal model. The idea is to find the seasonal pattern from the part of the series that most accurately reflects the seasonal pattern. To decide if your series needs to be shortened for identifying the seasonal model examine the data collection methods and the properties associated with the series.

Users of statistics, particularly economists, appreciate long time series. If you find notable inconsistencies in the series, provide a shorter seasonally adjusted series only and the longer historical series with sufficient quality warnings. Another option in treating inconsistencies is to provide two separate time series, one for the latest period and one for an earlier period, compiled with different methods or definitions.

For greater consistency, the changes in concepts, definitions and methodology may require corrections in the original time series. Statistical offices are in a better position to improve the quality and length of their time series than the users of statistics. The users tend to link new and old time series one way or the other if the statistical office doesn't do so.

Index calculation

The index calculation methodology affects the quality of time series, and thus, the results of seasonal adjustment. This area is well developed in the national statistical offices. However, in some countries, the focus has been on the current data, and more attention needs to be paid on time series methodology. This section may also offer examples for explaining indices to the users of statistics.

The index form is used both for intertemporal comparisons and for comparison of development between countries (Balk, 2008). The name "index" comes from Latin and means a pointer (UNSD, 2010). An index is a ratio that indicates the increase or decrease of a magnitude (Allen, 1975). The aim of an index, is to allow for comparison in time and space. In the economic field, indices help identify the different phases of the business cycle and monitor both short-term and long-term movements.

Indices serve different purposes. For example, a volume index measures the proportionate changes in the quantities of goods or services consumed or produced. A price index measures changes over time in the general level of prices of goods and services.

Usually, an index is described as a percentage of a base value that equals 100. It compares other observations to this base period. To consider a simple example where only one product is concerned, a cup of coffee may cost 3 Swiss francs in 2011, whereas it was 2.4 francs in 2005.

$$\frac{\textit{new observation}}{\textit{observation of the base period}} \Rightarrow \frac{3.0}{2.4} \times 100 = 125$$

In that example, the year 2005 is our base period. The index number for 2005 is 100, and the comparable figure for 2011 is 125. The result indicates a price increase of 25 per cent. In other words, you need to know today's price, quantity or value for a product and the equivalent for the base period. Today's observation represents the numerator and the base period the denominator. To get an index, the result is multiplied by one hundred.

The individual observations of an index series should refer to the same index reference period, for example to year 2010, which equals 100. Index observations referring to different reference periods, for example to the years 2005 and 2010, need to be linked to construct a series. With a common reference period observations can be directly compared. In some instances, statistical releases focus on the monthly (or quarterly) index that refers to the previous month or the same month of the previous year without offering any time series. For such figures, in which the index reference period changes from month to month, we can't perform seasonal adjustment.

In general, an index comprises three basic elements: prices, quantities and values. By knowing two of these elements one can calculate the third.

$$Price \times Quantity = Value$$

In calculating indices, the basic question is how to aggregate or summarize the possibly millions of prices and quantities into a smaller number of variables. To do this, we have to use weights. Traditionally, volume indices use a fixed-weight approach with weights updated, for instance, at five-yearly intervals. However, in recent years, chain-linking with annually updated weights has become a more preferred option, as it improves the timeliness of the applied weights.

The three most common index formulas are called by the names of their developers: Etienne Laspeyres, Hermann Paasche and Irving Fisher[3].

The Laspeyres-type volume index formula multiplies the quantity of produced goods at a given month by the prices at a base period. The value of production at the current period, based on the prices of the base period, is divided by the value of production at the base period.

$$I_L = \frac{\sum P_{i0} Q_{it}}{\sum P_{i0} Q_{i0}} \times 100, \textit{ where}$$

Pio = Price of product i at base period 0
Qio = Quantity of product i at base period 0
Qit = Quantity of product i at calculation time t

The Paasche-type index formula multiplies the quantity of a product at a given month by the price at the current period. The value of production at the time of calculation is divided by the value of production in prices of the current period and quantities of the base period.

$$I_P = \frac{\sum P_{it} Q_{it}}{\sum P_{it} Q_{i0}} \times 100, \textit{ where}$$

Pit = Price of product i at calculation time t

The Fisher index formula calculates the geometric mean of the Laspeyres and Paasche indices.

[3] For further detail on index calculation see: Balk, 2008.

Fixed weight indices have a weight structure selected at a point in time. The weights represent the relative importance of different products or activities at the weight reference period. Weights are used to compute indices over an extended period. The index compares the values always to the fixed base period. The weights of quantity indices are typically updated every five years (sometimes more frequently), and at the same time ideally the series is re-calculated based on the new weights.

Chain-linked indices refer to updating weights and linking the series with new and old weights together to produce a continuous time series. This approach doesn't include re-calculation of the entire historical series, but it rather links or splices together the two series. This way the chained index takes into account the changes in weighting structures for the duration of the series. *The International Recommendations for the Index of Industrial Production* (UNSD, 2010) provide practical instructions on chain-linking in chapter 5.

Cumulative indices are a measure of summarizing the development of the current year compared with the same period of the previous year. A cumulative figure may provide data for example for a period from January to May. In the next release, data could cover the period from January to June. Thus, the length of the reference period changes with each release of data, and the user cannot match revisions to the correct month. Therefore, the user cannot derive an accurate, discrete time series. Cumulative indices are useful as additional information, not as the main release form.

For an industrial production index, the index reference period should be a year. Weights could be updated at least every five years. For a chain-linked index the weights are updated more often, for example, annually. The more frequently the weights are updated the more representative will the resulting series be. (UNSD, 2010.)

Chain-linking has also a practical advantage in that it makes it easier to introduce new products or enterprises into the sample. It could improve consistency between structural and short-term statistics as a side result of updating the weights more often. The drawback is that the chain-linked totals and their components normally loose their consistency, also referred to as "additivity". Chain-linking requires timely data for the frequent weight updates which may also introduce structural changes into the time series.

We'll now look into those chain-linking practices that may influence seasonal adjustment. The lost additivity between components should be kept in mind in calculating the seasonally adjusted aggregates, because the aggregates adjusted individually could differ from the sum of their indirectly adjusted components. We'll discuss these aggregation techniques in chapter 5.

The alternative chain-linking techniques cause different types of breaks into the raw time series. Some of these breaks may have an impact on the seasonal pattern.

The three basic methods for chain-linking are: the annual overlap, one-quarter (or month) overlap and over-the-year approach. The weights in these approaches are expressed in the prices of the previous year. In the annual overlap approach the quantities refer to the average of the previous year; in the one-quarter overlap to the fourth quarter of the previous year; and for the over-the-year approach to the same quarter of the previous year.

A Eurostat/ECB Task Force on seasonal adjustment of quarterly national accounts has reviewed the impact of the chain-linking methods on seasonal adjustment. It concludes that the over-the-year technique interferes with the seasonality of the series more than the other methods and is not recommended. It may also hamper detection of outliers and identifying ARIMA models. The other two methods produce undistorted development within the calendar year, and the annual overlap method results in data consistent with a direct annual index.

The Task Force recommends for quarterly national accounts the use of the annual-overlap technique, widely applied by statistical offices and suggests performing seasonal adjustment after chain-linking of the original series to avoid the risk of artificial seasonality. On the other hand, price statistics commonly apply the one-quarter or month overlap. Statistical offices are to choose the most appropriate method for their statistics.

Measuring change consistently

The methods used in measuring change greatly influence the quality of the original time series, particularly its consistency which is important for seasonal adjustment. This section addresses the main issues which influence the consistency of time series.

The basis for producing a consistent time series is to have a documented definition of the statistical indicator. It has been said that statistical production is about bringing the data to the level of the intended statistical output (Eurostat, 2006). Before this can be done, the concept of the intended output has to be clear. The concept definition links the indicator to the world it's trying to measure.

Consensus on what and how should be measured makes compilation easier and improves coherence in time. If there's no common definition users will have varying perceptions about what's being measured, and the rules of data editing may change from one period to the other. The concept definition helps us understand the purpose of a statistical indicator, e.g. by documenting differences in definitions, scope and methods compared with other statistics.

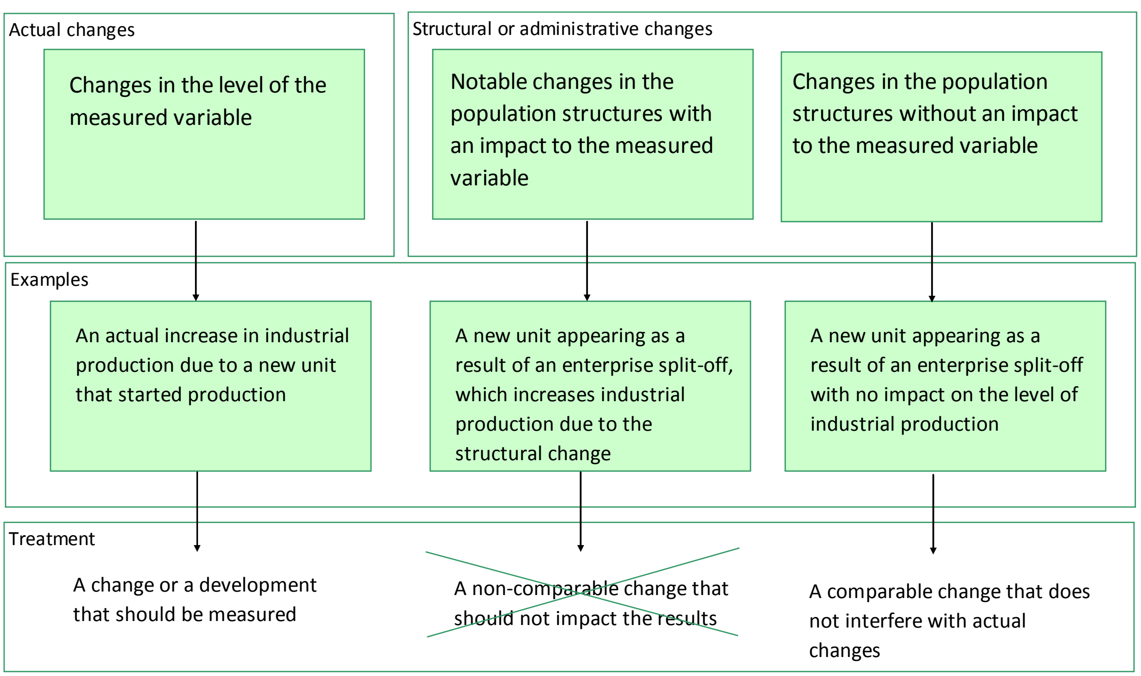
Part of the concept definition, is the way change is measured. In contrast to annual structural statistics, monthly and quarterly statistics don't aim to describe the level of turnover or industrial output but rather their development in the course of time. In general, the structural statistics describe the variable at a given point in time without attempting to make different periods of time comparable with each other. But the main problem in compiling short-term statistics is how to achieve comparable measures of development to build a consistent time series.

Questionnaire design should reflect the concept definition of the statistical indicator and give the necessary instructions and examples to the respondents. To enable better consistency in time, respondents should be allowed to report revisions to earlier data on the questionnaire. Pre-filled questionnaires increase accuracy of source data, and at the same time, accuracy of seasonality, as the respondents may check and revise the previously reported monthly or quarterly values.

The population covered by monthly or quarterly statistics doesn't consist of exactly the same units at different reference periods. For example, an enterprise may change the type of activity and end up in a different industrial activity class. New enterprises may start, while others may close down. Enterprises may change their structure through takeovers, mergers or split-offs.

Changes in the population don't always result from actual changes in the measured variable, such as the level of economic activity. The core question in measuring changes is the different treatment of actual changes and non-comparable changes. Actual changes that affect economic activity should influence a time series directly; whereas non-comparable changes shouldn't influence the results (Eurostat, 2006). The only change that should

Chart 1
Types of changes taking place in business population



affect a time series is the measured development itself, i.e. an actual change (chart 1).

For industrial production index, the only changes that should influence the index are changes in industrial production itself, and not, for example, changes in enterprise structures.

Structural changes in the population may include changes in the relations between units. Sometimes these are linked to each other and shouldn't be treated separately. If the linked changes cancel one another, and thus don't significantly influence the results, they make up a comparable change.

New units or new products may enter the population at any time. Ideally they should appear in the index compilation on entering the market. The same logic applies to unit closures and exits. If a unit stops doing the measured activity, it should have a decreasing effect on the index.

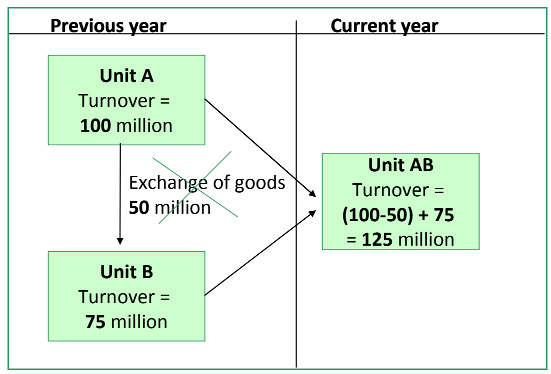
However, some new units may appear as a result of enterprise split-offs. These units could be recognizable by their size which exceeds the normal size of a newly set up enterprise. If a new unit is created by a structural change and would affect the index, the effect shouldn't be included in calculation. Large, yet genuine, new units may appear, exceptionally for example, when international chains enter new countries. The entering of new units should be visible in the series if they describe a real change in the economic activity, even if only temporary.

An existing unit may be reclassified out of the measured activity. A unit could either have changed its activities radically, for example, by having moved from industry to services or the change may have been a gradual one. If the change actually took place around the time of the reclassification it should affect the results. However, pure administrative correction of an activity class or gradual changes that finally lead to a new activity classification shouldn't greatly influence the results.

If a closure happens as a result of an enterprise merger, and if the partners of the merger don't add up approximately to the same level before and after the merger, the effect shouldn't be taken into account. Some partners of a merger may not operate in the measured activity class, in which case the figures may not match. Also, if the partners of a merger are engaged in bilateral trade, the merger may decrease the level of activity (chart 2, Eurostat, 2006). This is due to the trade between the merged units no longer being reported, having become an internal activity of the unit. Such a problem applies specifically to statistics compiled at the enterprise level.

Data collection could include questions designed for identifying non-comparable changes or confirming actual changes. In case of large percentual changes in an observation, the respondent could be asked to give additional information to reply, for example, if there have been shop openings or closings, enterprise mergers or split-offs. However, a test group comprising some of the respondents should check any changes to questionnaire design. This kind of survey testing could reveal surprising issues the survey maker should be aware of.

Chart 2
The impact of a merger on comparability



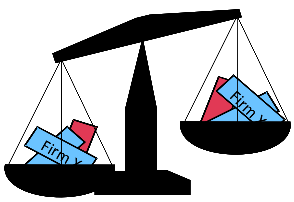
Producing consistent time series requires, in practice, good compilation tools to avoid excessive use of time and manual work. The quality of raw data can be improved by defining systematic data editing rules; the idea being not to check all observations but to introduce plausibility checks to quickly find those observations that need special attention or correction.

Efficient data editing draws attention to the errors that have a significant influence on the results, to values that are outside a given range of variation and to those that are not coherent compared with the related data sources. Information from other sources, such as the business register or registers of other government offices could help confirm unit closures and structural changes.

A traditional check is to compare the new observations to past values, e.g. to the same month of the previous year and then pay attention to those observations which deviate significantly from their past behaviour. These rules of thumb help detect any significant non-comparable changes.

Dedicate some time to large, atypical observations. Careful consideration of outliers during data editing makes seasonal adjustment easier. Atypical values may also be correct but may be caused by unusual circumstances. If so, they should remain visible in the data. If the atypical value turns out to be a mistake, it should be corrected by finding out the real value or by estimating.

There are three main approaches to treating changes in the population. Firstly, to record all changes the way they are. The benefit of this approach is simplicity, but the resulting statistics are difficult to interpret. As these changes don't reflect only the measured activity, it may become challenging to obtain a correct picture of the development in the course of time.



The second option is to use a pure panel method. This includes only the comparable observations of the reference period and the base period. Each observation is only included if it has a comparable pair. This method, too, is simple but may lead to biased results in dynamic populations because it doesn't include unit start-ups or closures, as there is no comparable data for them for both periods.

The perhaps ideal, yet complex, method, sometimes called an overlapping method (Eurostat, 2006), allows only comparable changes in the measured activity to influence the results. The method aims at correcting errors and non-comparable changes by collecting data for the comparable units, estimating a comparable observation or taking the non-comparable unit out of the calculation scheme. This approach would reflect actual changes in the measured activity, but it requires much more work. As a result of measuring only actual changes of the variable, time series would be more consistent; and administrative and structural changes wouldn't distort the results.

The theoretical aim is to get results that would arrive if the structures at the reference and base period were equal. To compile indices without non-comparable changes, you have to modify the results either for the base or for the reporting period. This causes some practical complexities.

In the backward-oriented method, all units have the same structure they had in the base period. The forward-oriented method modifies all previous periods to make them comparable with the current one.

Other compilation issues also influence the consistency of time series, but we do not discuss them here in more detail. These could include sample design, treatment of non-response and methods of using administrative or other additional data sources.

Components of time series

Main components

Seasonal adjustment divides a time series into its different parts to identify patterns of the series that can be modelled and forecasted. As described in chapter 1, these include a seasonal, irregular and trend-cycle component. Applying seasonal adjustment, in other words, means extracting the seasonal component from the raw series.

Chart 3 introduces the main components of a time series. The original series, on the left, contains all the characteristics of a time series without any adjustments. Next to it is the seasonal component which should be periodical and repeat itself almost the same way from one year to the next.

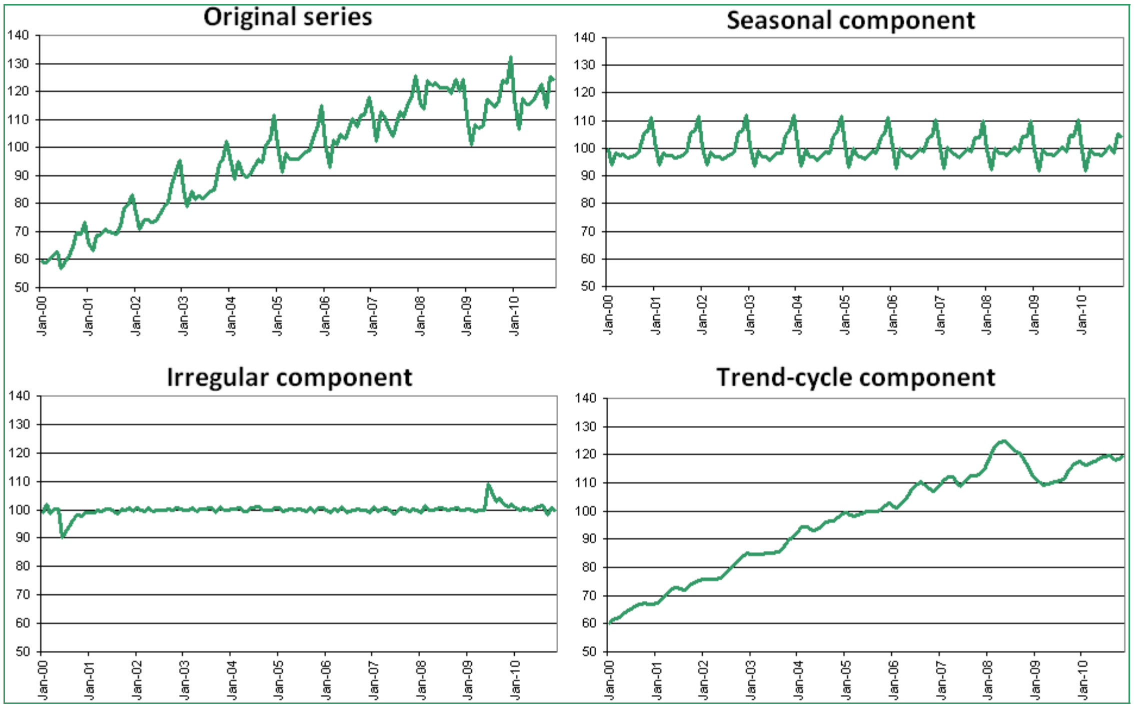
Since the seasonal component may hide the underlying movement of the series, TRAMO/SEATS or X-12-ARIMA extracts it from the data. If the series is affected by calendar effects these are removed before applying seasonal adjustment.

The seasonally adjusted series is, thus, the result of deducting the seasonal component, including the calendar effects, from the original data. At the same time, the seasonally adjusted series is a combination of the trend-cycle and the irregular component.

The seasonally adjusted series contains the irregular component which includes noise and some outliers. These sudden events could be strikes or some other extreme changes in the data. As the irregular component includes outliers, these will also remain visible in the seasonally adjusted data. By extracting the trend-cycle component from the seasonally adjusted, you end up with the irregular component. The trend-cycle series is relatively smooth and evolves over the long-term without much disturbance.

Chart 3

Components of Time Series of the Industrial Production Index of Kazakhstan



Original series contains all characteristics of the series, without any adjustments or reduction. Both random and systematic fluctuations influence the series. It may contain seasonal effects and effects related to the calendar, such as moving holidays. In other words, it doesn't make two months comparable with regard to the amount of working or trading days. Its random variations include the irregular fluctuations and extreme values.

Calendar effects refer to any economic effect related to the calendar, and these are removed before seasonal adjustment and decomposition of a series. The leap year effect, the number of different days appearing in a month, the composition of working and non-working days as well as moving holidays may alter the level of activity measured at a point of time.

Seasonal component refers to those fluctuations, observed during the year, which repeat on a fairly regular basis from one year to another. The timing, direction and magnitude is more or less stable. The seasonal component includes seasonal and calendar related effects. The seasonal effects may be caused by social events and habits as well as seasonal changes in the weather. The calendar effects refer to the differing numbers of trading days and moving holidays in a month or quarter.

Irregular component captures the remaining short-term fluctuations in the series which are neither systematic nor predictable. It is assumed to include only white noise. The irregular component is the remaining component after the seasonal and trend-cycle components have been removed from the original data. It may also contain some errors which are not necessarily random.

Transitory component is sometimes identified and estimated by SEATS. While the irregular component is forced to be white noise, the transitory component picks up short-term (transitory) variation other than that. It would interfere with estimating the seasonal component. Once the components are estimated, the irregular and transitory component can be put together. For simplification, we'll discuss them together.

Trend-cycle component includes both long-term and medium-term developments. Out of these two, the trend depicts the long-term evolution over several decades, i.e. structural changes. It reflects the underlying level of the series and is typically the result of population growth, price inflation, technological change and general economic development. The cycle component, on the other hand, is the relatively smooth movement around the long-term trend from expansion to recession.[4]

[4] See a course: www.cros-portal.eu/page/estp-courses

Decomposition type

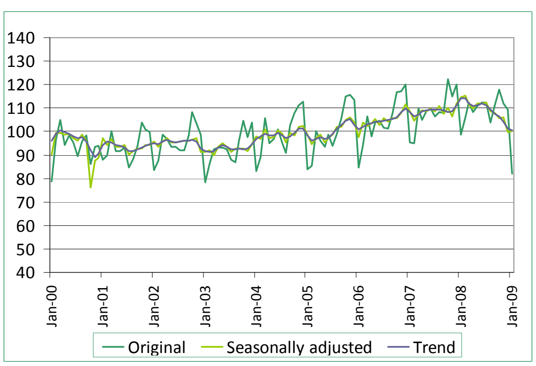
Seasonal adjustment estimates and identifies the different components of the original time series. Decomposition assumes that the components of the series behave independently of each other. The two main types of decomposition are "additive" and "multiplicative". TRAMO and Reg-ARIMA automatically identify the most suitable model based on the TRAMO procedure.

For additive decomposition (chart 4), the magnitude of seasonal effects doesn't change as the level of the trend-cycle changes. Also, any series with zero or negative values are additive. The additive decomposition means that the time series adds up as a sum of its independent components: trend-cycle + seasonal + irregular(/transitory[5]).

$$\textit{Additive Model} : X_t = TC_t + S_t + I_t$$

Here X_t stands for the observed value of the time series at time t. The seasonal component is S_t for the given point in time. The trend-cycle component is TC_t and the irregular component I_t. The difference between a cyclical and a seasonal component is that the seasonal one occurs at regular intervals annually, while cyclical factors have a longer duration and vary around the trend.

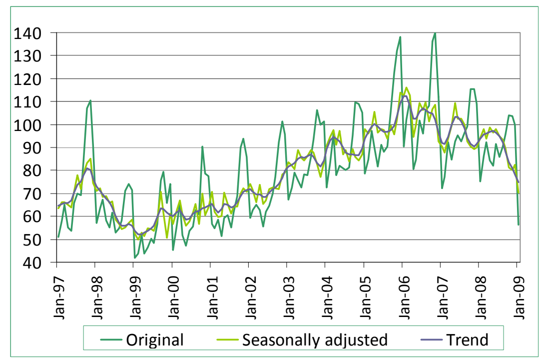
Chart 4
An example of an additive decomposition



The multiplicative decomposition implies that as the trend of the series increases, the magnitude of the seasonal spikes also increases (chart 5). For a multiplicative decomposition, the original time series is the product of its components: trend-cycle x seasonal x irregular(/transitory). Most economic time series exhibit a multiplicative relation between the components. If the magnitude of the seasonal effects increases when the level of the series increases and vice versa, and the series has no negative values, a multiplicative decomposition can be applied.

$$\textit{Multiplicative Model} : X_t = TC_t \times S_t \times I_t$$

Chart 5
An example of a multiplicative decomposition



Effects influencing time series

Seasonality

In general, finding seasonality in a series is a precondition for performing seasonal adjustment, but seasonality shouldn't be present any more in the adjusted data. Paying attention to seasonality stems from the fact that the seasons exert a notable influence on economic and social activity.

The fluctuations in a time series are either repeatable or non-repeatable. By definition, *seasonality* is a pattern of a time series in which the data experience regular and somewhat predictable changes that repeat every calendar year. Seasonality refers to periodic fluctuations, e.g. the tendency of retail sales to peak during the Christmas season. Seasonality is quite common in economic time series.

Human activity includes rhythms such as the 24-hour rotation of days and social habits, the alternation between night and day, the weekly rest, holiday periods and consumption habits that vary according to the season. Seasonality reflects traditional behaviour associated with the calendar. It may also include the impact of business habits, such as quarterly provisional tax payments or periodic invoicing and administrative procedures such as the timing of tax returns. Any

5 For example, the models (1-0.3 B)x_t= a_t or (1+0.3 B^{12})x_t = a_t cannot generate a trend-cycle, have no seasonal component, and are not white noise. They will generate a transitory component.

administrative or accounting rhythms visible in the data could mislead the interpretation of statistics.

People's habits change from summer to winter, for example because of the differences in the temperature. These changes are captured by the seasonal component. Extreme weather conditions, however, would be part of the irregular component, as they don't repeat every year.

Without seasonal adjustment we could say, for example, that unemployment decreases in the summer and rises again in autumn. However, this is merely usual fluctuation, as many summer workers are hired to take the place of regular personnel during their summer holidays. Seasonal fluctuations in data make it difficult to analyse whether changes are actually important increases or decreases in the level of activity, or if they are part of the regular variation.

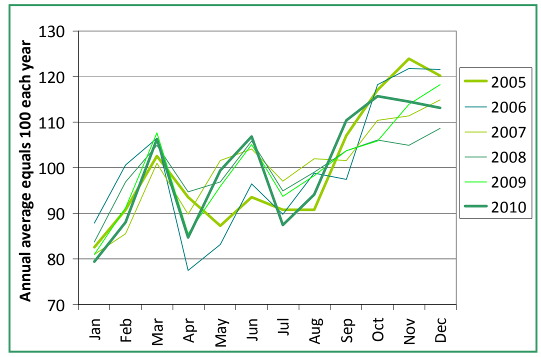
Seasonality is the pattern that TRAMO/SEATS or X-12-ARIMA extracts from the original series in order to produce the seasonally adjusted series. Working day correction is an additional procedure that helps identify the underlying movement. It aims at obtaining a series whose values are independent of the length and the composition in days of a month (Eurostat, 2009). Moving holidays also cause problems for interpretation of data. Adjustment of moving holidays estimates and corrects for those effects differing across years.

We can usually detect seasonality by doing a visual analysis of the data. By comparing the annual original series of different years, we should be able to observe some of the repeating patterns.

The example of an industrial production index of the Republic of Moldova displays a clear seasonal pattern (chart 6) in the original series. The series grows towards the end of the year, and a smaller peak repeats, for instance, in March. Although the pattern is repetitive, there's some variation in the level of the monthly observations because the original series includes not only the seasonal component but also the trend-cycle and the irregular variations. For example, the index for October varies between 104 and 124 points depending on the year, although the annual average of each year has been set to 100.

The seasonal component may include some moving seasonality. In practice, seasonality usually evolves over time, and the seasonal factors reflect this.

Chart 6
Monthly fluctuation of the industrial production index of the Republic of Moldova, original series



Calendar effects

The seasonal component contains different types of influences: the seasonal effect and calendar effects, including moving holidays. These effects may influence the level of activity in a month or quarter. The varying weekdays and the leap year effect are called calendar effects (Findley and Soukup, 2000). These effects include the composition of weekdays and variation of non-working and working days as well as moving holidays. The leap year influences the time series, as it adds one day in every four years, and that day can be either a working day or it may fall into the weekend.

Months have an excess of different types of days. If the level of an activity is higher on some days, for example on weekdays, then the series may have a working day effect. In many industries, for example, the volume of production is smaller on Saturdays and Sundays than on weekdays. Thus, the level of activity in a given month is likely to be higher if the month contains more weekdays and lower if it contains more weekend days than on average.

Some calendar effects are seasonal and are thus part of the seasonal component. For example, the length of months is a seasonal phenomenon. For this reason, the calendar component only contains the part of the non-seasonal calendar effects.

National holidays influence the number of working days. Since they vary from country to country, a country specific historical list of public holidays and compensation holidays, if any, would be useful for estimating the calendar component. Compensation holidays refer to additional holidays sometimes announced if a national holiday falls into the weekend. Moving holidays could be particularly country specific due the timing of different religious holidays that varies between countries. Some religious holidays, such as Easter and Ramadan, occur each year, but with a different date. The most common moving holiday is Easter which moves between March and April.

The fact that Easter moves between months may have a marked influence on the level of activity (table 4). The table illustrates the importance of working day correction. For example, April 2008 had three more working days than April 2007. This was because in 2008 Easter was not in April, whereas in the previous year it was. By just comparing April 2008 to April 2007, without any adjustment, industrial production seems to have grown about 8 per cent. However, after working day adjustment, this growth fell to 5 per cent.

Table 4
The impact of moving Easter effect on growth rates for April 2008

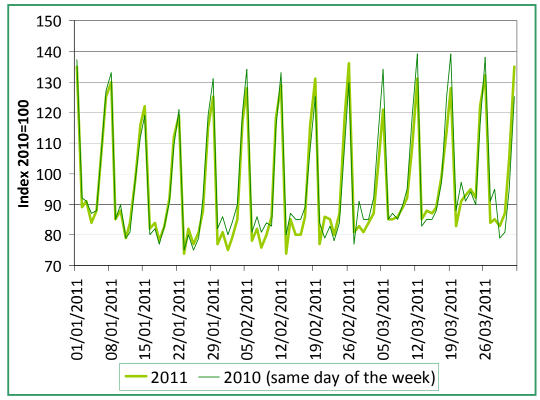
Growth rates (April 2008)		
	from April 2007	**from March 2008**
Original series	**8.1**	-
Working day adjusted	**5.1**	-
Seasonally adjusted	-	**1.2**

The seasonally adjusted data bring another angle to the development. With the seasonally adjusted data, the growth rate can be calculated from the previous month, as seasonal adjustment makes all months comparable with each other. Judging by the seasonally adjusted data (table 4), industrial production rose by 1 per cent in April 2008, compared with March 2008. After all, it seems that the industrial production was only merely growing. The working day adjusted growth, 5 per cent, provides useful information but is calculated in comparison with a period one year before (April 2007). In spring 2008, the growth seems to have levelled out.

Institutional arrangements in different industries lead to visible differences in the number of working days in a particular industry. Sometimes, the number of weekdays doesn't affect the level of activity at all, for example, when the factories operate in shifts every day.

A real data example illustrates the effect that different weekdays, sometimes also called trading days, can have (chart 7). The curves depict the daily sales of an individual shop in two years, in 2007 and 2008. Based on the varying sales one sees here that Saturday is a very good day for business. With this kind of fluctuation, the number of Saturdays in a month affects directly the amount of sales in a given month: the more Saturdays there are, the more sales.

Chart 7
The sales recorded on different trading days



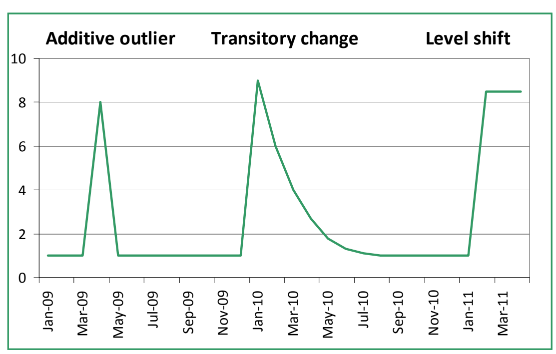
Outliers and interruptions

Most time series contain some volatility, causing the original and seasonally adjusted values to fluctuate around the trend. Time series usually contain some extreme values that deviate from the trend by a large margin and fall outside the range expected based on the typical pattern of the time series. These extreme values are called "outiers".

Some adverse natural events or industrial disputes may cause outliers. The adoption of a new law or a new form of subsidy, as well as a new type of tax, could also cause them.

There are at least three kinds of outliers as mentioned in chapter 1 (see chart 8). An additive outlier is a single point jump in the time series; a transitory change is a point jump followed by a smooth return to the original path; and a level shift is a more permanent change in the level of the

Chart 8
The most common types of outliers



series. Some other types of outliers, such as seasonal outliers, could be added to the list. Those refer to breaks in the seasonal pattern. They aren't currently supported by Demetra+. Since the outliers include information about specific, unusual events, such as strikes, they are useful as part of the seasonally adjusted data, and they remain visible.

An additive outlier is an impulse outlier that affects only the value of one observation. It could be caused by random effects, strikes or bad weather. A pre-announced price rise could cause an additive outlier by increasing the sales dramatically before the price change. Additive outliers need to be identified since seasonal adjustment is based on moving averages, and is therefore, sensitive to extreme values.

For *a transitory change*, gradual reducing or rising follows the extremely high or low observation until the time series returns to the initial level. Transitory changes could occur due to deviations from average monthly weather conditions. If in the winter the weather suddenly becomes colder than usually the energy consumption would probably go up. When the weather gradually returns to the average level the consumption should settle back to normal.

A level shift refers to a change into higher or lower values starting from a given time period. This means that the level of the time series undergoes a more permanent change. Sometimes a change in concepts and definitions or compilation methods of a survey may cause this kind of a shift. But statisticians should try to maintain the comparability of the series without causing a level shift because of a methodological change. Level shifts often occur as a result of changed economic behaviour, new legislation or changed social traditions. For example, if the salaries increase for some profession, the level of that time series becomes permanently higher, but the seasonal pattern does not change. These need to be identified so that they don't distort the seasonal component.

Outliers in the latest month or quarter are difficult to identify. Unfortunately, real extraordinary economic effects are often unknown, and there's no information on what happens after the latest outlier appears. Before seeing more observations, you can't distinguish a level shift from an additive outlier since we don't know how the level of the series will behave. Knowing the series and having external information of the event in question would help define the type of outlier. Conduct a continuous analysis to identify reasons for outliers, and where possible, document the explanations for them.

A large number of outliers in a time series may make it difficult to identify a seasonal model. This may signify a problem with the reliability of the data. A large number of outliers relative to the length of the series could result in over-specification of the regression model. In the production phase, a statistician needs to check all outliers and correct possible errors before seasonal adjustment.

Time series may also contain other interruptions. For example, seasonal breaks could occur especially in highly volatile time series. This means that the seasonal pattern could change suddenly to a different pattern that starts to repeat in the series thereafter. Seasonal breaks are often accompanied by a level shift. The cause could be a structural change because of a crisis or a methodological or an administrative change, such as timing of invoicing.

Missing observations could also interrupt the series. Series with too many missing values will cause estimation problems in the adjustment. Thus, statisticians should substitute the missing observations with alternative data or statistical methods in the lack of original data.

Requirements for input data

Demetra+ has specific requirements for the format of input data, even though it provides flexible alternatives for reading data that are in different formats. The different formats supported by Demetra+ include Excel, databases through Open Database Connectivity (ODBC), Statistical Data and Metadata Exchange (SDMX), WEB services and several types of files such as text, TRAMO/SEATS for Windows (TSW), X-12-ARIMA software by the U.S. Census Bureau (USCB), XML etc. The access is dynamic: the software automatically refreshes time series by consulting the original source file or database to download new information.

ODBC is software for accessing database systems independently of programming languages, database and operating systems. Thus, any application can use ODBC to extract data from a database.

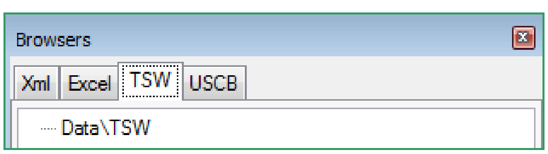
The next chapters will make use of example data sets that include industrial production indices (2005=100) of the countries of Eastern Europe, Caucasus, Central Asia and China[6]. Table 5 shows the properties of these example datasets for the purposes of this analysis. As you can see, the length and frequency of data differ between countries. All series in the dataset are technically appropriate for performing seasonal adjustment since they meet the minimum length criteria for time series.

Table 5
Basic features of the example dataset

Countries	Data Covered		Period	Number of Obs.
	Start	**End**		
Armenia	Jan-2000	Nov-2010	Monthly	131
Azerbaijan	Jan-1996	Oct-2010	Monthly	178
Belarus	Jan-1997	Nov-2010	Monthly	167
China	Jan-2000	Nov-2010	Monthly	131
Georgia	Q1-2001	Q3-2010	Quarterly	39
Kazakhstan	Jan-2000	Nov-2010	Monthly	131
Kyrgyzstan	Jan-1993	Oct-2010	Monthly	214
Moldova, Republic of	Jan-1997	Oct-2010	Monthly	166
Tajikistan	Jan-1997	Oct-2010	Monthly	166
Ukraine	Jan-2000	Oct-2010	Monthly	130

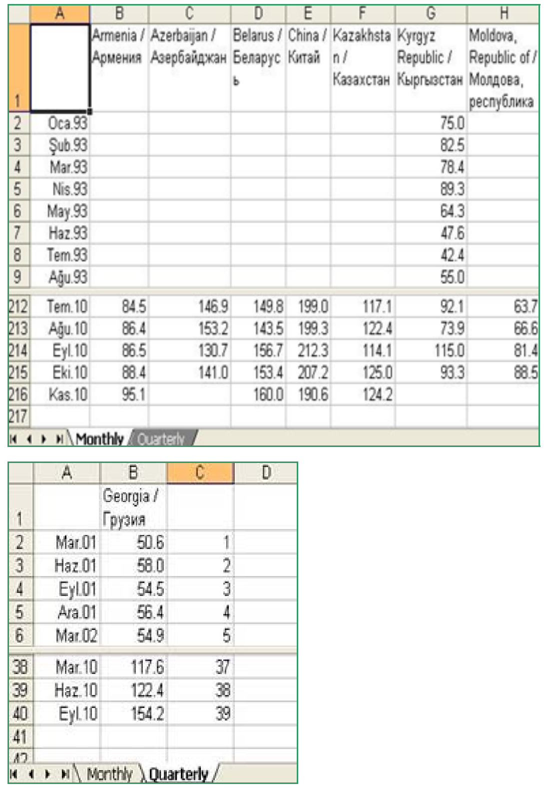
The *Browsers* panel lists the most usual source file types on the left side of Demetra+ (image 30). To open the source file, **click on the correct sheet**, such as TSW and **locate the file**. The data will appear in the navigation tree of the TSW provider. The set of time series in a sub-folder will be grouped under a collection called All series. The same idea applies for the USCB and XML files.

Image 30
File types supported by Demetra+



Demetra+ sets some rules to the layout of the input data. Image 31 shows an example of a vertical presentation of input data; but a horizontal presentation is also possible, and up to transposition of the cells, follows exactly the same rules.

Image 31
An example of the structure of input data for monthly and quarterly time series



	A	B	C	D	E	F	G	H
1		Armenia / Армения	Azerbaijan / Азербайджан	Belarus / Беларусь	China / Китай	Kazakhstan / Казахстан	Kyrgyz Republic / Кыргызстан	Moldova, Republic of / Молдова, республика
2	Oca.93						75.0	
3	Şub.93						82.5	
4	Mar.93						78.4	
5	Nis.93						89.3	
6	May.93						64.3	
7	Haz.93						47.6	
8	Tem.93						42.4	
9	Ağu.93						55.0	
212	Tem.10	84.5	146.9	149.8	199.0	117.1	92.1	63.7
213	Ağu.10	86.4	153.2	143.5	199.3	122.4	73.9	66.6
214	Eyl.10	86.5	130.7	156.7	212.3	114.1	115.0	81.4
215	Eki.10	88.4	141.0	153.4	207.2	125.0	93.3	88.5
216	Kas.10	95.1		160.0	190.6	124.2		
217								

Monthly / Quarterly

	A	B	C	D
1		Georgia / Грузия		
2	Mar.01	50.6	1	
3	Haz.01	58.0	2	
4	Eyl.01	54.5	3	
5	Ara.01	56.4	4	
6	Mar.02	54.9	5	
38	Mar.10	117.6	37	
39	Haz.10	122.4	38	
40	Eyl.10	154.2	39	
41				

Monthly / Quarterly

Demetra+ requires that in all sheets, the cell A1 must be empty and the column A must contain the reference dates of the observations. The cells of this column need to use true dates as can be defined in the Excel. The date may refer to any day during the reference month or quarter. The first row, i.e. the cells B1 and C1 etc. should contain the titles of each series. Inside a sheet, the titles have to be unique. The cells below each title contain the observations of the time series.

Missing data are indicated simply by an empty cell. Demetra+ identifies the series in a workbook both by the name of the sheet that contains them and by the title of each series. You shouldn't change this core information if you wish to update the analysis in a second adjustment dynamically based on the same workbook.

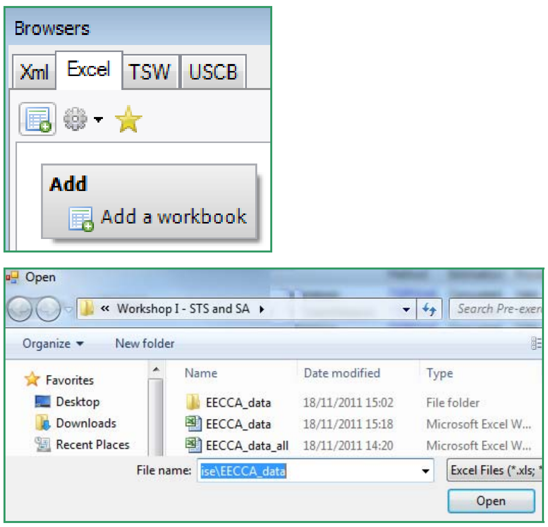
As indicated in table 5, these instructions will make use of nine monthly and one quarterly series. If you have data on multiple frequencies, you should divide them into two sheets. Therefore, the name of the first sheet to be used in this analysis is Monthly, and the second sheet is called Quarterly.

[6] The countries of Eastern Europe, Caucasus and Central Asia with available data were covered; data were not available for Turkmenistan and Uzbekistan. In addition, the industrial production index of China is included in the example.

Once the data are in the described format, Demetra+ is able to receive them. You can import data directly from an Excel file by the clipboard or by dragging and dropping. However, time series imported this way can't be updated automatically.

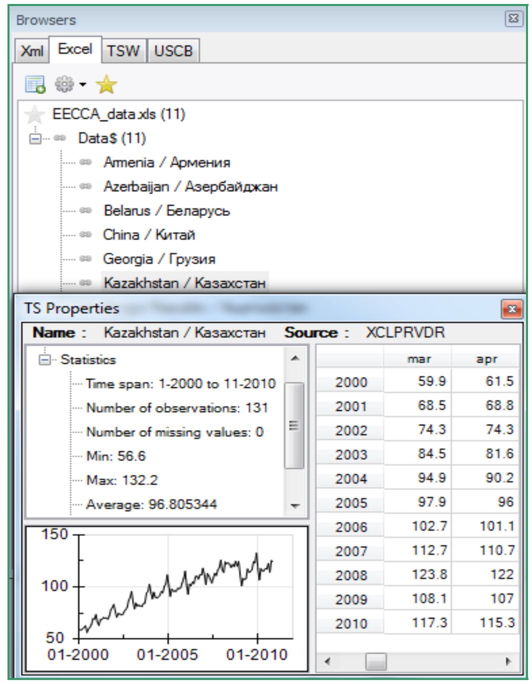
To enable repetitive use of the source file, you may prefer to load the dataset into Demetra+ with tools provided by the software itself. To do this, go to the *Browsers* panel, **click on the Excel tab** and then click on the **Add** button to add an Excel workbook (image 32). The program opens a standard dialog window. **Select the Excel file** and **press Open**, the name of the Excel file and the number of the series (in brackets) will appear in the *Browsers* panel.

Image 32
Importing an Excel file



The tree in the *Browsers* panel also shows the different sheets of the Excel file and the series in each sheet. By **clicking on one of the series** listed in the *Browsers*, for example, Kazakhstan, the series appears in the *TS Properties* panel. The panel displays summary statistics of the series: max, min, average, etc. It shows a graph and a table of the time series (image 33).

Image 33
Viewing the basic features of a time series in Demetra+



Visual checking of original series

To assess the prerequisites for seasonal adjustment, Demetra+ supplies a wide range of useful visual tools depending on the selected adjustment method. These tools help evaluate the attributes of the time series, the magnitude of seasonal variations, presence of seasonality and breaks in the seasonality or trend.

The human eye can observe remarkable breaks or interference in the series. Nevertheless, some atypical behaviour may be part of real development. Rarely does any measured variable vary according to a pre-defined model. Visual analysis offers a quick tool for statisticians to identify the possible outliers, missing values or errors. However, the statisticians should identify and correct errors before seasonal adjustment, during editing of data and compilation of indices.

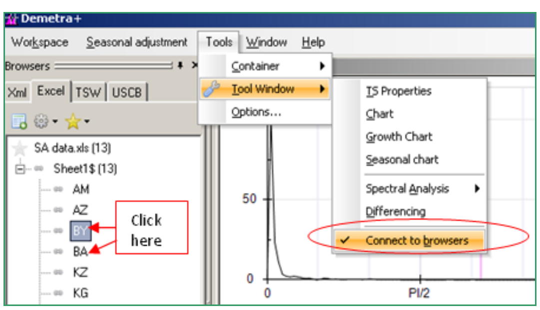
The visual graphs provide the user with quick identification of possible problems and quality issues. Once identified, the reasons for these atypical observations or sudden changes in the pattern of the time series need to be investigated. Good documentation of the weaknesses of the raw data prepared during the compilation phase would be helpful. Not only is good documentation necessary for the producer, but it also benefits the user of statistics. If there's no documentation to explain the observations made during visual analysis of the raw data, the correctness of the data should be checked before seasonal adjustment.

To look at the time series in a table format, select **Tools/Container/Grid**. First, the Grid is empty and ready to collect one or more time series by **drag and drop**, you can grab the name of the time series with the mouse from the *Browsers* panel and drop it in the Grid.

Demetra+ calls the different visualisation tools containers. For a closer visual analysis, you can select **Tools/Container/Chart**. You can add a time series to a container by **dragging and dropping** it from the *Browsers* panel or from the Grid which shows the data in a table form. You can include several time series in the same graph and remove them one by one by selecting them in the graph. Other containers, such as Chart and Growth chart are available under the same menu. All these containers become active by dragging and dropping data from the *Browsers* or from the Grid.

If you're working with a set of time series and would like to check them visually, first open a tool from **Tools/Tool Window**, then **Connect to Browsers** panel (image 34). This way the Chart or Grid will update every time you **click on a name of a series** in the *Browsers* panel. This provides a fast tool for visual checking of multiple time series one after the other.

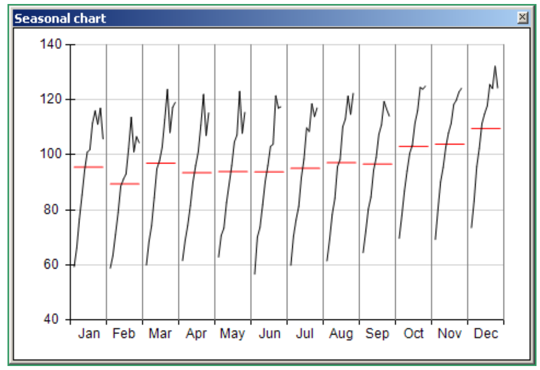
Image 34
Visual checking of multiple series



The Tool Window of Demetra+ offers some graphs specifically designed for analysing seasonality of a time series. Seasonal graphs visualize how months or quarters differ from each other and how much the observations for the same month differ between the years. To obtain a Seasonal chart, select **Tools/Tool Window/Seasonal Chart** and drag and drop the series into the window.

A seasonal graph is a special form of line graph that displays all the observations for one month or one quarter successively (image 35). In other words, it gives first all the observations for each January of the time series. Horizontal lines on the chart illustrate the average of each period, e.g. the average of all Januaries.

Image 35
Seasonal graph of the industrial production index of Kazakhstan

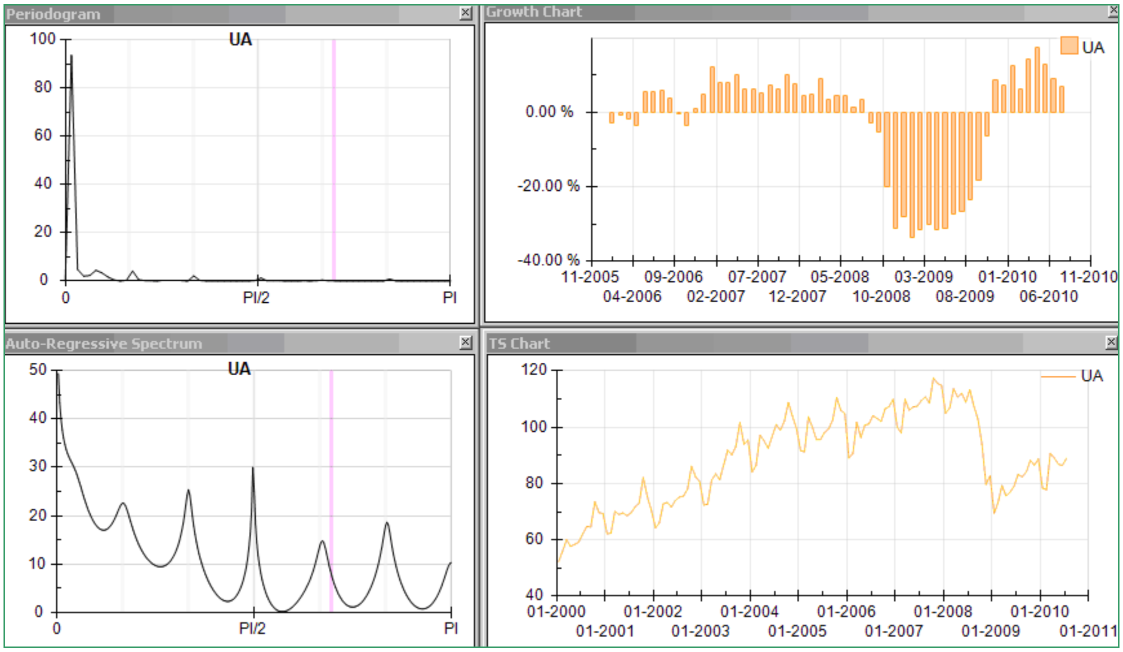


The chart shows that Kazakhstan's industrial production is typically at its lowest value in February, and it peaks in December. The differences between the monthly averages reflect the non-comparability of the original data for different months. For example, in Kazakhstan industrial production is likely to be higher in March than in February each year. Comparing March and February as such doesn't give any information about the underlying development.

Statisticians tend to use spectral analysis, the Periodogram and the Auto-Regressive Spectrum, for confirming that seasonality is present in the original series or for detecting remaining seasonality after adjustment. To visualize the chart, choose **Tools/Spectral Analysis/Periodogram**.

In the graph (image 36), judging by the peaks at the grey vertical lines, the auto-regressive spectrum (bottom, left) displays clear seasonality in the original data. It shows peaks at the zero-frequency which indicate the presence of a trend-cycle component. It doesn't show any distortion at the frequency related to trading day effects (the purple line). These kinds of plots are useful also for detecting unwanted, remaining seasonality after seasonal adjustment.

Image 36
Periodogram and the auto-regressive spectrum

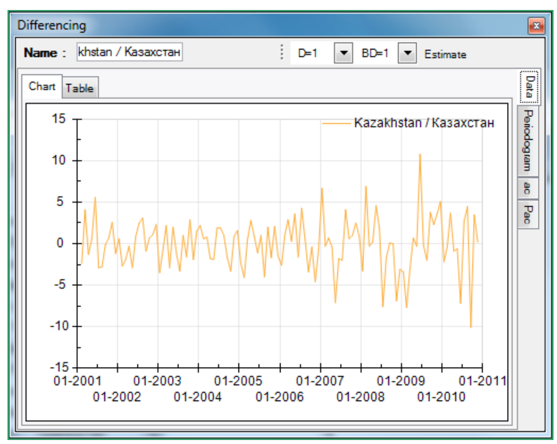


For further analysis, Demetra+ also includes a Differencing tool. It can be used for defining the ARIMA model by determining the order of differencing. You can open it by selecting **Tools/Tool Window/Differencing**. You can drag and drop the time series into the name box of the window. In general, the term differencing refers to making a time series stationary, so that it would be easier to predict.

The estimated regular (D) and seasonal differences (BD) are displayed on top of the Differencing tool and can be changed manually. The sub-windows of the tool display the differenced series, its graph, its periodogram and auto-correlation (image 37).

Based on the graph, the differenced series is stationary, i.e. its statistical properties such as mean, variance and autocorrelation are relatively constant over time. As a conclusion, one can expect that the ARIMA model of both TRAMO/SEATS and X-12-ARIMA would identify similar results in the pre-adjustment of data.

Image 37
Results of the differencing tool for the industrial production index of Kazakhstan



CHAPTER 3

Seasonal adjustment phase

Introduction

This chapter will now look at the second phase, i.e. the actual seasonal adjustment, using industrial production indices as an illustrative example. First, we'll examine the important issue of defining and using national calendars in Demetra+. Using national calendars improves the estimate of the calendar effects. Some national calendars are available in the software, but others are added or imported in a format suitable for Demetra+. After that, we'll look at how to seasonally adjust one time series, i.e. apply the single processing. In the later part of this chapter, we'll set out instructions for defining and launching multi-processing, i.e. seasonal adjustment of a set of series simultaneously.

Currently, the most popular seasonal adjustment approaches are X-12-ARIMA, developed by the US Census Bureau[7], and TRAMO/SEATS, developed by the Bank of Spain[8]. To bring these two methods closer, Eurostat developed the Demetra+ software which provides statisticians with an interface to both methods.

Sometimes the seasonally adjusted series and the seasonal component are called the "unobserved components", as they cannot be directly extracted from a time series. Statisticians have developed ways for estimating these series. Historically, analysts often used a technique that consisted of drawing, by hand, the seasonally adjusted series based on the visual inspection of the raw series. The coming of computers enabled applying seasonal adjustment with mathematical tools and treatment of a great number of time series efficiently.

One of the first mathematical methods for estimating the trend applies moving averages. In this method, the monthly average of the last 12 months forms the observation for each period. After withdrawing the trend-cycle, based on moving averages, from the original series, the result will comprise the sum of the seasonal and irregular component. This method, however, has some weaknesses. For example, the cyclical fluctuation could be allotted, by error, with the seasonal component. And outliers influence the estimation of the trend-cycle. The problem is also that the moving average over twelve months necessarily leads to a six-month time lag in estimation, as it's derived based on past observations. The end of the trend-cycle series is uncertain and new observations have a significant effect on the results.

Nonetheless, moving averages still form the basis of seasonal adjustment methods. Taking into account the weaknesses of moving averages, new methods using weighted averages were developed. As a result, in 1954, the United States Census Bureau introduced a new method, called Method I (Bell and Hillmer, 1984), for computer assisted seasonal adjustment, which was followed by Method II. The later versions of this method were named X for experimental and called X-0, X-1, X-2... until the widely used X-11 model was introduced in the 1960's. The X-11 performs seasonal adjustment by carrying out several iterations to smoothen the series with moving averages.

The X-11 applies moving averages to arrive at the trend-cycle and seasonal component. The raw series is divided by the seasonal component in order to obtain the seasonally adjusted series in case of multiplicative decomposition. In case of an additive composition, the seasonal component is deducted from the raw series. This method still has many of the weaknesses of the simple moving averages. To overcome the time delay caused by the average of the previous 12 months, one could either increase the weight of the uncertain last observations or stretch the time series at both ends by forecasting. X-11-ARIMA method was developed to rectify this weakness.

The X-11-ARIMA, introduced by Statistics Canada, adds estimates at the beginning and at the end of the series before proceeding to seasonal adjustment. These estimates are done by forecasting and backcasting with the ARIMA part.

7 For more information, see: www.census.gov/srd/www/x12a/

8 For more information, see: www.bde.es > services > statistics > statistics and econometrics programs

The forecasted ends of the series significantly decrease the size of revisions and the time lag of estimation. This improves the precision of the seasonal component. The AR of ARIMA comes from the word auto-regressive and means that the value is determined by a relation with the previous values of the series. The MA of ARIMA comes from moving average part of the ARIMA model, and finally, the I means integrated.

The US Census Bureau developed the X-11-ARIMA further to the method called X-12-ARIMA (Findley et all, 1998). As well as applying the basic idea of its predecessor, X-12-ARIMA offers flexible tools for pre-adjustment (Reg-ARIMA), detects and corrects for calendar effects and outliers by means of a regression-ARIMA type model before identifying the seasonal component. X-12-ARIMA offers a comprehensive set of diagnostic tools important for checking the quality of the results (ECB, 2000).

The other approach for seasonal adjustment is signal extraction, used by TRAMO/SEATS. This approach is based on describing the behaviour of the series, thus, deriving the different components of the series as captured by an ARIMA model. This leads to consistency between the ARIMA model obtained for the observed series and the models of the components. Victor Gómez and Agustín Maravall of the Bank of Spain developed these two programs for applied time series analysis. You may turn to Gómez and Maravall (2001) for a comprehensive description of the software, and to Maravall and Pérez (2011) for information on the latest facilities and interpretation of the SEATS results.

In TRAMO/SEATS, the first part, TRAMO, is responsible for pre-adjustment. TRAMO refers to Time Series Regression with Arima Noise, Missing Observations and Outliers. It is a program for (often automatic) identification, estimation, forecasting and interpolation of Regression-ARIMA models. This model is used to extend the series with forecasts and to identify and estimate outliers, calendar and other effects by regression variables. You may use the software on its own for outlier detection or forecasting, for example, but it's mainly used for pre-adjustment before actual seasonal adjustment. TRAMO has many similarities with the pre-adjustment tool of the current X-12-ARIMA.

In the second step, TRAMO passes the series forward to SEATS which applies a filter to the extended series net of outliers and regression effects. This way SEATS obtains estimators of the seasonally adjusted series and the seasonal, trend-cycle and irregular component. By combining them with the outliers and regression effects, the final estimators of the components are obtained. SEATS is an abbreviation from Signal Extraction in ARIMA Time Series. It's a program for decomposing a time series into its components, following an ARIMA-model-based method. In fact, the starting point for SEATS was a preliminary program built for seasonal adjustment at the Bank of England, in 1982 (Gómez & Maravall, 1996).

The methods, X-12-ARIMA and TRAMO/SEATS have common features. They first perform pre-adjustment which adjusts for the working day effect and outliers by means of a regression model. Secondly, they identify and estimate the trend-cycle, seasonal and irregular components from the time series and sometimes also a transitory component. The two methods are very similar in the first part but differ in the adjustment step. The filter of X-12-ARIMA is selected from a set of pre-designed filters, whilst SEATS follows a so-called ARIMA-model based approach. Furthermore, their outputs and quality diagnostics take different forms.

It has been said that the combination of these two approaches in one seasonal adjustment tool would be a promising way forward (see ECB, 2000). Demetra+ now brings these two methods closer to each other by offering them under the same interface. With Demetra+ it's easy to compare the differences of the seasonally adjusted series between these methods. We'll now take a look at seasonal adjustment with Demetra+ and some practical examples.

Define calendars

Demetra+ provides an easy tool for creating calendar regression variables using the Calendar module. The calendar regression variable can include trading or working day variables as well as country or sector specific holidays. These holidays may occur at fixed or moving dates.

If the national calendar doesn't include extraordinarily complex moving holidays, you can use the pre-defined options of Demetra+ for modelling it. If it includes several moving holidays that start or were abolished during the time series, you could set a validity period for the holidays or prepare the national calendar in a format of an external user-defined variable, i.e. as a time series to be imported to Demetra+. In the next section we'll discuss first the pre-defined options.

In Demetra+, TRAMO/SEATS and X-12-ARIMA automatically create appropriate trading day, working day, leap year and Easter regression variables depending on the chosen specifications. However, the user may need to change the automatic options for:

- Modifying the trading day, working day and leap year regression variables to match the national holidays that differ from the pre-specified options of Demetra+.
- Chaining two calendars for two different time periods.
- Combining two or more calendar variables using proper weights.
- Importing user-regression variables including holidays with changing frequency, duration, start and end dates.

Demetra+ offers the following pre-defined regression variables:

- Six variables to test for trading day effects: *(N Mondays) – (N Sundays), (N Tuesdays) – (N Sundays),…, (N Saturdays) – (N Sundays)* where N means the number of.
- One variable to test the effect of weekdays versus weekend days.[9]
- Leap year variable.
- Easter variable.

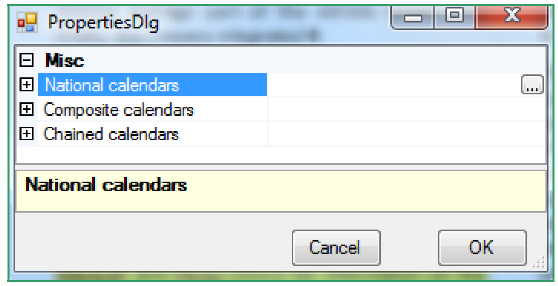
The leap year variable measures the effect of differences in the number of days in Februaries. This variable equals zero for all months different from February. In February, it takes the value -0.25 if the February comprises 28 days and -0.75 for 29 days.

TRAMO/SEATS assumes that the Easter effect lasts for six days. For X-12-ARIMA, the RSA5(c) specification tests for three different lengths of the Easter effect and selects the optimal length from 1, 8 or 15 days. The user may also define the duration. The variable includes a zero for all months different from March and April since Easter occurs in March or in April (Gómez and Maravall, 1996).

You may define the calendars from the Main menu, **Workspace/Edit/Calendars**. This module creates new calendar regression variables which incorporate both country-specific holidays and default calendar regression variables.

The Calendar module offers a possibility to define National Calendars, Composite Calendars or Chained Calendars (image 38).

Image 38
Different calendar options available in Demetra+

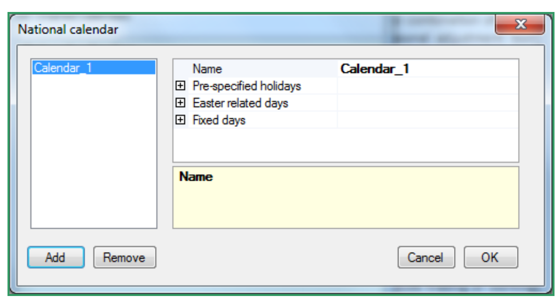


To add a new calendar, activate the row National calendars and **click on the** **button** and **Add**. The options for defining the national calendars will appear. These include Pre-specified holidays, Easter holidays and Fixed days (image 39).

Pre-specified holidays

Demetra+ includes 13 regression variables for modelling pre-specified holidays based on the most common European holidays. To select national holidays, activate the row **Pre-specified holidays** and **click on the** **button** next to it (image 39). Once you add a holiday from the list, you may also add the offset and the validity period. We'll explain these terms in the following example.

Image 39
Options for national calendars



[9] The variable may be defined as *(N(M,T,W,Th, F))-(N(Sat,Sun)×5/2* (Gómez and Maravall, 1996)

As an example, the following holidays could be defined for Belgium (table 6). They include nine holidays available under the pre-specified holidays. The National day, Assumption of Mary and Armistice Day are not available under the pre-specified holidays.

Table 6
List of holidays in the example of Belgium

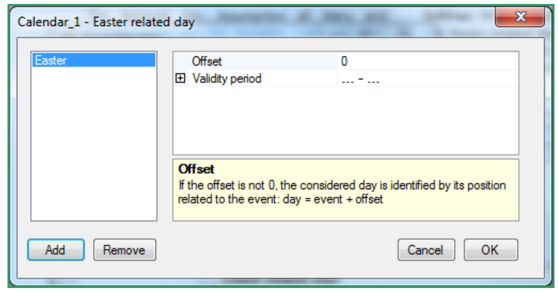
Holiday
New Year's Day
Easter
Easter Monday
Labour Day
Ascension
Pentecost
Pentecost Monday (WhitMonday)
National day
Assumption of Mary
All Saints
Armistice Day
Christmas

Demetra+ aggregates all the holidays added by the user into the regression variable. In table 7 you can find the definitions of the pre-specified holidays available in Demetra+.

Easter-related holidays

The duration and the exact timing of Easter and Easter-related days may vary in different countries. The Easter-related days include Pentecost and Whit Monday. Also, the number of holidays around Easter may vary. Therefore, Demetra+ provides a feature called Easter-related days for the user to design exactly the country-specific **Easter-related holidays**. When you **click on the** ... **button** next to the Easter-related days, you can **Add** the related holidays (image 40). The relationship between Easter and the related days is determined by offset.

Image 40
Selecting Easter-related days



As the Easter-related holidays vary between countries, let's assume that Pentecost and Whit Monday wouldn't be available in the pre-specified holidays. These holidays occur as defined in table 7.

The calendar regression variable for Belgium was created in the previous example of pre-specified holidays. For exercise, Pentecost and Whit Monday will be entered as Easter-related days. Therefore, you need to remove them from the selected pre-specified holidays. The number of pre-specified holidays for Belgium is now seven.

Table 7
Definitions of the pre-specified holidays in Demetra+

Holidays	General Definitions
New Year	falls on 1 January
AshWednesday	a moving holiday, occurring 46 days before Easter
Easter	a moving holiday; it varies between 22 March and 25 April depending on offset time
MaundyThursday	falls on the Thursday before Easter. It is the fifth day of Holy Week, and is preceded by Holy Wednesday and followed by Good Friday
GoodFriday	refers to the Friday in Easter week
EasterMonday	the day after Easter Sunday
Ascension Day	traditionally celebrated on a Thursday, the 40^{th} day from Easter day
Pentecost	celebrated seven weeks (50 days) after Easter Sunday, Pentecost falls on the tenth day after Ascension Thursday
WhitMonday	the holiday celebrated the day after Pentecost, a movable feast in the Christian calendar. It is movable because it is determined by the date of Easter
MayDay	on 1 May is a spring festival and usually a public holiday
Halloween	an annual holiday observed on 31 October
AllSaintsDay	a solemnity celebrated on 1 November by parts of Western Christianity, and on the first Sunday after Pentecost in Eastern Christianity
ThanksGiving	currently, in Canada, Thanksgiving is celebrated on the second Monday of October and in the United States, it is celebrated on the fourth Thursday of November
Christmas Day	a holiday observed on 25 December in most countries

Click on the **button** next to the **Easter-related days**. Since the Pentecost holiday occurs 50 days after Easter the offset is set to 50. If it were to occur 50 days before Easter the user would enter 50 with a minus (-) sign. For Whit Monday, the offset value will be 51 since the holiday is celebrated one day after Pentecost.

Fixed holidays

Countries usually also have fixed holidays in their national holiday calendar. These holidays always occur on the same date. You may add these exact dates by **clicking on the** **button** next to the **Fixed days** and **enter the dates**. Sometimes the fixed holidays may occur during the weekend and sometimes during the week. This information can't be given to Demetra+ by using the currently available tools for setting fixed holidays.

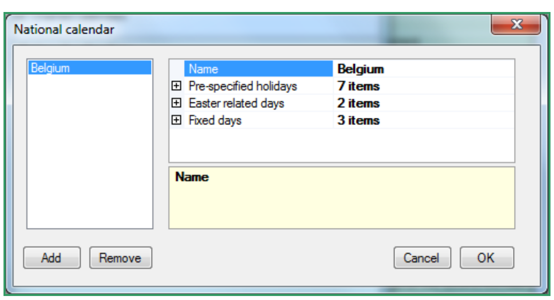
The National day, Assumption of Mary and Armistice Day are not available for the pre-specified holidays, but they occur on fixed dates every year (table 8) and can be added as fixed holidays.

Table 8
Fixed holidays in the Belgian example

Holiday	Date
National holiday	July 21
Assumption of Mary	August 15
Armistice Day	November 11

The resulting full holiday calendar for Belgium, used in this example, is illustrated in image 41. It includes seven pre-specified holidays, two Easter-related holidays and three fixed dates that repeat each year.

Image 41
The example holiday calendar for Belgium



Composite calendars and chained calendars

Demetra+ offers an option to define a composite calendar. Sometimes it could be useful, e.g. for preparing a calendar for the aggregate industrial production index of the EU-countries. With this application, the users can combine the calendar regression variables previously created by defining appropriate weights for each calendar.

On a rare occasion, the user of Demetra+ may need to chain two different calendar regression variables. This request may arise for a holiday being removed which existed before a certain date. You can chain calendar regression variables by selecting the calendars and a break date.

After selecting **Chained calendar** instead of a National calendar (as seen earlier in image 38), **click on the** **button** next to the **Chained calendars**, and you can **Add** the required calendars and the break date. You should first add the calendar to be used in the first part of the chained regression variable.

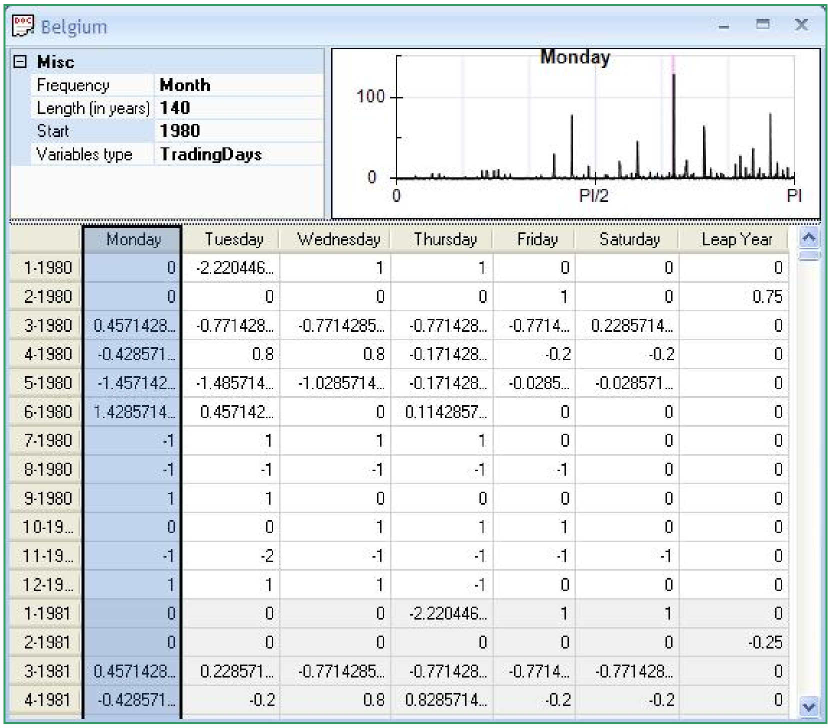
Viewing the calendar regression variables

Demetra+ presents the calendar regression variable created for Belgium by **a double click on the calendar** called "Belgium" on the *Workspace* panel. Image 42 displays the regression variable. The upper-left panel, Misc, displays the properties of the calendar regression variable. The upper-right panel presents the selected regression variable graphically.

The Misc panel allows the user to access the frequency of the regression variable, length in years, start date and variable type. The user can transform the regression variable: the option None of the variable type leads to Demetra+ applying only the leap year variable. The option Trading days refers to the six trading day variables and the leap year variable.

The last option Working days represents the calendar regression variable which models the difference of activity between the weekend and the weekdays and the leap year variable (image 42). The regression variables generated can be copied by drag and drop into Excel.

Image 42
Viewing the calendar regression variables in Demetra+



	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Leap Year
1-1980	0	-2.220446...	1	1	0	0	0
2-1980	0	0	0	0	1	0	0.75
3-1980	0.4571428...	-0.771428...	-0.7714285...	-0.771428...	-0.7714...	0.2285714...	0
4-1980	-0.428571...	0.8	0.8	-0.171428...	-0.2	-0.2	0
5-1980	-1.457142...	-1.485714...	-1.0285714...	-0.171428...	-0.0285...	-0.028571...	0
6-1980	1.4285714...	0.457142...	0	0.1142857...	0	0	0
7-1980	-1	1	1	1	0	0	0
8-1980	-1	-1	-1	-1	-1	0	0
9-1980	1	1	0	0	0	0	0
10-19...	0	0	1	1	1	0	0
11-19...	-1	-2	-1	-1	-1	-1	0
12-19...	1	1	1	-1	0	0	0
1-1981	0	0	0	-2.220446...	1	1	0
2-1981	0	0	0	0	0	0	-0.25
3-1981	0.4571428...	0.228571...	-0.7714285...	-0.771428...	-0.7714...	-0.771428...	0
4-1981	-0.428571...	-0.2	0.8	0.8285714...	-0.2	-0.2	0

External user-defined variables

The user may want to add specific calendar regression variables by defining them outside of Demetra+. Some moving holidays such as Ramadan or the Feast of Sacrifice may require using an external user-defined variable.

To do this, you could prepare the regression variables in Excel and import them to Demetra+ as a user-defined variable. This section shows how to create such an external variable.

First, the dates of Ramadan and the Feast of Sacrifice should be determined day-by-day for a specific length. In this example, the required length is from 1 January 1974 to 31 December 2015.

Table 9 displays the design of the required Excel file. The columns from A to F represent the official holidays in Turkey. Saturdays and Sundays are also listed in the file to see when they coincide with the dates of Ramadan or the Feast of Sacrifice.

For each moving holiday value 1 is given when it occurs on the specific date. This is measured in two ways: first for each time the holiday occurs excluding Saturdays, Sundays and official holidays, and second for each time the holiday occurs excluding Sundays and official holidays.

This list is then constructed for the required period of time for which the regression variable is needed.

Table 9
Design of an Excel file for compiling an external calendar regression variable

			A	B	C	D	E	F			G	H	J	K
Date	Sacrifice	Ramadan	New Year	23 April	1May	19 May	30 Agu	29 Oct	Saturday	Sunday	Sacrifice=1 ; otherwise=0 (Excl. Sat, Sun and official holidays)	Sacrifice=1 ; otherwise=0 (Excl. Sun and official holidays)	Ramadan=1 ; otherwise=0 (Excl. Sat, Sun and official holidays)	Ramadan=1 ; otherwise=0 (Excl. Sun and official holidays)
1 Jan 1974 Tue	0	0	**1**	0	0	0	0	0	0	0	0	0	0	0
2 Jan 1974 Wed	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3 Jan 1974 Thu	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4 Jan 1974 Fri	**1**	0	0	0	0	0	0	0	0	0	**1**	**1**	0	0
5 Jan 1974 Sat	**1**	0	0	0	0	0	0	0	**1**	0	0	**1**	0	0
6 Jan 1974 Sun	**1**	0	0	0	0	0	0	0	0	**1**	0	0	0	0
7 Jan 1974 Mon	**1**	0	0	0	0	0	0	0	0	0	**1**	**1**	0	0
8 Jan 1974 Tue	0	0	0	0	0	0	0	0	0	0	0	0	0	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
26 Dec 2015 Sat	0	0	0	0	0	0	0	0	**1**	0	0	0	0	0
27 Dec 2015 Sun	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0
28 Dec 2015 Mon	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29 Dec 2015 Tue	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30 Dec 2015 Wed	0	0	0	0	0	0	0	0	0	0	0	0	0	0
31 Dec 2015 Thu	0	0	0	0	0	0	0	0	0	0	0	0	0	0

As shown in table 10, the second step is to aggregate the results for each month.

Table 10
Aggregated monthly values

Date	Sacrifice Holidays (Excl. Sat, Sun and official holidays)	Ramadan Holidays (Excl. Sat, Sun and official holidays)	Sacrifice Holidays (Excl. Sun and official holidays)	Ramadan Holidays (Excl. Sun and official holidays)
Jan.74	2	0	3	0
Feb.74	0	0	0	0
Mar.74	0	0	0	0
Apr.74	0	0	0	0
May.74	0	0	0	0
Jun.74	0	0	0	0
⋮	⋮	⋮	⋮	⋮
Jul.15	0	**1**	0	2
Agu.15	0	0	0	0
Sep.15	3	0	**4**	0
Oct.15	0	0	0	0
Nov.15	0	0	0	0
Dec.15	0	0	0	0

The third step is to extract the variables from the averages of each month (table 11). Demetra+ assumes that a user-defined regression variable is appropriately centered, i.e. the mean of each user-defined regression variable is subtracted from the means for each calendar period. One has to be careful with the centered regression variable, as changes in the number of public holidays will influence the monthly average for the whole time span.

The calendar variable should be centered, first, not to cause any level effect on the series. Second, the monthly average should be extracted from the calendar variable since the calendar variable may contain seasonality. The calendar variable should measure only calendar effects, not seasonality.

The average value for the regression variable is calculated for each month. The user can directly deduct the sum of the Ramadan and Sacrifice holiday regression variables from the monthly average to obtain a total holiday regression variable.

Table 11
Feast of Sacrifice holiday regression variable

Date	Sacrifice Holidays (Excl. Sat, Sun and official holidays)	Average values for 1974-2015	**Sacrifice Regressor (Excl. Sat, Sun and official holidays)**
Jan.74	2	0.2381	**1.7619**
Feb.74	0	0.2143	**-0.2143**
Mar.74	0	0.2143	**-0.2143**
Apr.74	0	0.1905	**-0.1905**
May.74	0	0.1429	**-0.1429**
Jun.74	0	0.2143	**-0.2143**
⋮	⋮	⋮	⋮
Agu.15	0	0.2143	**-0.2143**
Sep.15	3	0.2619	**2.7381**
Oct.15	0	0.3333	**-0.3333**
Nov.15	0	0.4048	**-0.4048**
Dec.15	0	0.2857	**-0.2857**

To import the user-defined calendar regression variables, you may drag and drop them into Demetra+ to make a static variable or directly read them from an Excel file to make dynamic variables. For the second option, select the Main menu and **Workspace/User variables**, or **double click on the User-defined variables** in the *Workspace* panel and drag and drop the time series from Excel with the key combination **ctrl-c** and **ctrl-v**. To apply this imported regression variable, you need to change the type of specifications for calendar effects to **UserDefined**.

Single processing

Select the approach

Demetra+ refers to the processing of one time series at a time as single processing. It offers several different processes for adjusting a single time series. First of all, you need to select the method from the two options: TRAMO/SEATS and X-12-ARIMA. You may try seasonal adjustment with both methods to see the difference.

Select specifications and regression variables

One of the first things to decide is the choice of specification that determines the way Demetra+ needs to proceed in seasonal adjustment. The pre-defined specification alternatives appear in the *Workspace* panel, and one could start experimenting first with the default specifications. To many users, these default specifications are flexible enough, but if need be, you may define additional specifications, as we'll explain later in this chapter.

The chosen specification tells Demetra+ whether TRAMO and Reg-ARIMA should test for the trading and working day effects as well as for the leap year effect etc. They perform the tests and apply the regression variables only if the effects are significant. If included in the specification, TRAMO and Reg-ARIMA will also test for the Easter effect, but in Demetra+ they don't test for all moving holidays by default as these vary between countries. In case of national moving holidays, one can import the national holiday calendar to Demetra+ as described earlier.

Demetra+ provides a set of pre-defined specifications, presented separately for TRAMO/SEATS and X-12-ARIMA. The options vary from the Airline model (RSA0) to more complex specifications (RSA5). Table 12 presents the alternative specifications for TRAMO/SEATS and X-12-ARIMA.

These default specifications appear in the *Workspace* panel, on the right side of Demetra+. The user should start their analysis with one of these specifications. Usually, they could start by choosing either the specification RSA4(c) or RSA5(c) and afterwards change some of the options, if need be. The difference between these is that RSA4 tests the difference between working days and non-working days, while RSA5 looks for a difference between the days of the week. The choice depends on the properties of the series. For example, if calendar effects are not significant in the series it should be reflected in the choice of specification.

If you want to frequently use a specification that isn't available in the predefined list, for example to integrate systematically your own calendar variables or to exclude some kinds of outliers, the best solution is to define your own specification and add it to the *Workspace*.

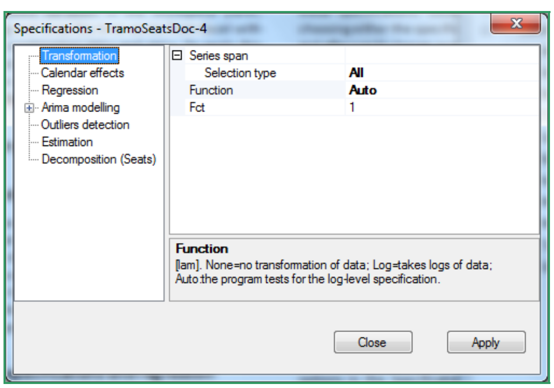
To add a specification, select **Seasonal Adjustment/Specifications** from the Main menu and **click on** TRAMO/SEATS **specification** or X-12-ARIMA specification. After choosing the needed options in the Specifications dialog box, the new specification will automatically appear in the *Workspace* panel. When you save the *Workspace*, it will include the new specification for future use. You may use any user-defined specifications the same was as the predefined specifications.

Table 12
Summary definitions of specifications presented in Demetra+

Method	Name	Explanation
TRAMO/SEATS	RSA0	level, Airline model
TRAMO/SEATS	RSA1	log/level, outlier detection, airline model
TRAMO/SEATS	RSA2	log/level, working days, Easter, outlier detection, airline model
TRAMO/SEATS	RSA3	log/level, outlier detection, automatic model identification
TRAMO/SEATS	RSA4	log/level, working days, Easter, outlier detection, automatic model identification
TRAMO/SEATS	RSA5	log/level, trading days, Easter, outlier detection, automatic model identification
X-12-ARIMA	X11	No pre-processing, only for decomposition
X-12-ARIMA	RSA1c	log/level, outlier detection, Airline model
X-12-ARIMA	RSA2c	log/level, working days, Easter, outlier detection, Airline model; pre-adjustment for leap year if log-transformation applied
X-12-ARIMA	RSA3c	log/level, outlier detection, automatic model identification
X-12-ARIMA	RSA4c	log/level, working days, Easter, outlier detection, automatic model identification; pre-adjustment for leap year if log-transformation applied
X-12-ARIMA	RSA5c	log/level, trading days, Easter, outlier detection, automatic model identification; pre-adjustment for leap year if log-transformation applied

The user may define the details of a specification by means of the Main menu for TRAMO/SEATS or X-12-ARIMA. Transformation page includes the definitions for the **time span** used for identifying the seasonal pattern. Most often, the series is analysed in its full length when the selection type is all and the auto function is chosen (image 43).

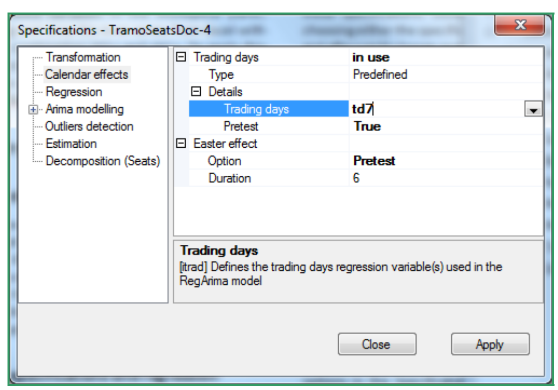
Image 43
Defining specifications, time span



If the pre-defined specifications aren't exact enough for the user's purposes, the **options for calendar adjustment** can be adjusted (image 44). The user can choose between no calendar adjustment, predefined type of adjustment, defined calendar or own calendar. In the next image, choosing None would mean that TRAMO won't analyse the presence of calendar effects at all. Predefined option would lead to using a default calendar.

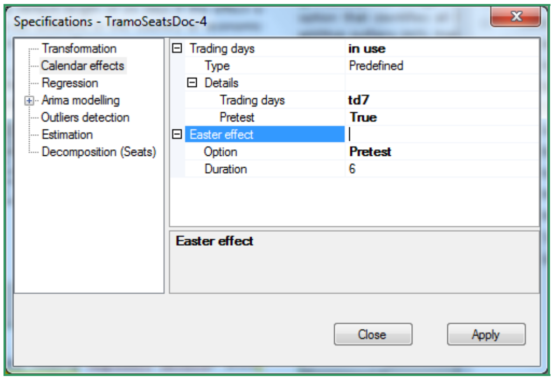
The option Calendar corresponds to the pre-defined trading day variables, modified to take into account the specific holidays according to users selections between td1, td2, td6 or td7. For TRAMO/SEATS, td1 includes the weekday-weekend contrast variable, where td2 adds the leap year effect to the working day effect. Td6 tests for the different effects of the days of the week and td7 adds to the test for the different days of the week the leap year effect. You may consult the *Demetra+ User Manual* for details for using X-12-ARIMA. You may select the User-defined option when you need to specify your own trading day variables.

Image 44
Defining specifications, trading days



One can also choose whether or not to apply a test for the **presence of Easter effect** (image 45). TRAMO will test for the need to correct for the Easter when Pretest is chosen, but the option Yes would automatically lead to the correction for Easter effect. The user can modify the length of this effect from the default length of six days (for TRAMO-SEATS) if the effect is likely to be different in the country or economic activity in question.

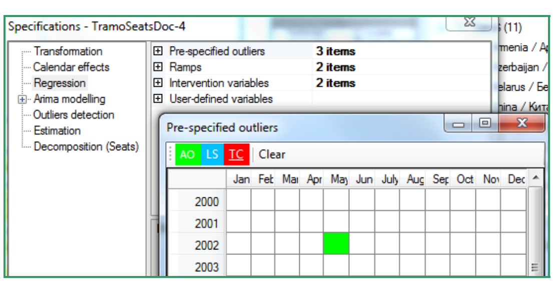
Image 45
Defining specifications, pre-test for Easter



Under the heading **Regressors** Demetra+ allows setting **dates of the outliers** that you're aware of before adjustment (image 46). These you may define if you have prior knowledge of such events and their date. However, the first adjustment of a series without preselected outliers reveals which outliers TRAMO or Reg-ARIMA would identify.

You may also set up ramps, intervention variables or user-defined variables under the specifications. Ramp effect means a linear increase or decrease in the level of a series that lasts for a certain period. The intervention variables refer to known special events, such as strikes, political decisions etc. which influence the data. It enables not only setting up outliers but also sophisticated interventions to match particular events.

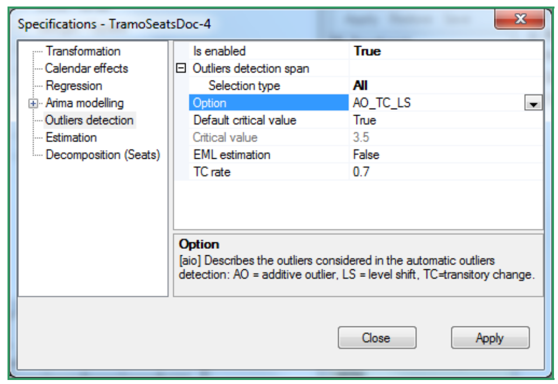
Image 46
Defining specifications, pre-specifying outliers



To enable automatic modelling during the seasonal adjustment, under the heading **Arima modelling/Automatic modelling** select **Is enabled/True**.

Enable **outlier detection** by choosing **Is enabled/True** (image 47). You could select the option that identifies all three types of outliers: additive outliers (AO), transitory changes (TC) and level shifts (LS). The default critical values provide best results for first seasonal adjustment, and modifying them requires experience. You may sometimes need to limit the time span of automatic outlier detection, for instance, to reduce revisions.

Image 47
Defining specifications, enabling outlier detection



For the remaining steps in defining specifications, namely Estimation and Decomposition, you may use the predefined specifications.

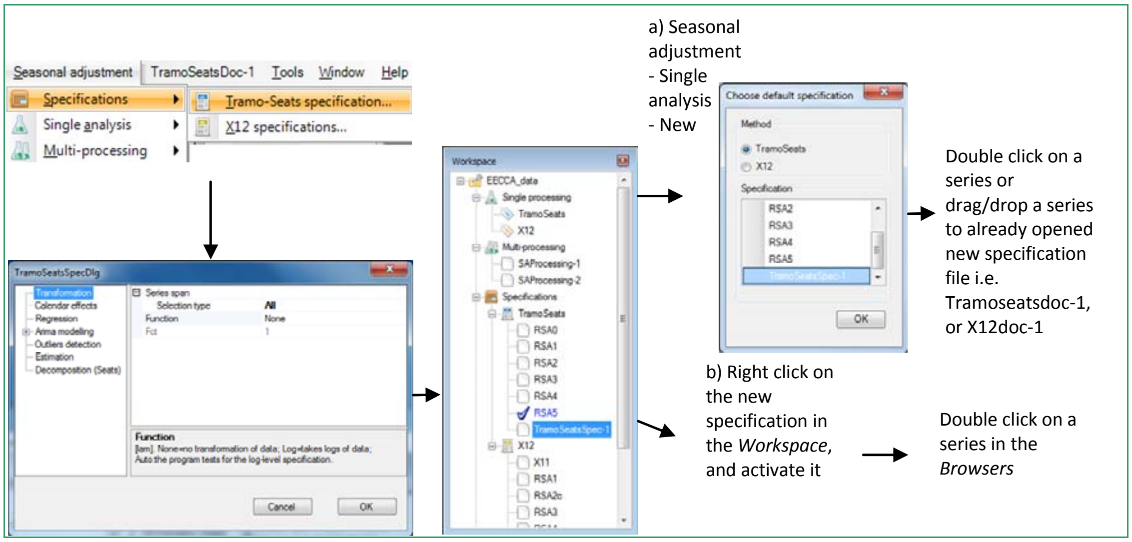
Launch seasonal adjustment

For adjusting a single time series, Demetra+ provides many alternatives from a simple double click on the name of the series to defining each step by using the *Wizard*.

First, you need to select an active specification. You may choose a specification by making a specification active **by right-clicking on** it in the *Workspace* panel and selecting **Active**. If other single processing windows are open, a new series will appear in all of them, even if they don't correspond to the active specification.

Once the specification is active, one of the options for launching an automatic single processing is to **double click on the name of a series** in the *Browsers* panel. If single processing windows aren't open, a new window appears. If you've already selected a specification in the current *Workspace*, Demetra+ applies this active specification.

Image 48
Creating a single-process with a user-defined specification



The second option to launch single processing is to **drag and drop the selected specification** to the *Results* panel from the *Workspace* panel. A new single processing window will open automatically. To import data into the window, either **double click on a series** in the *Browsers* or drag and drop the series into the left panel of the single processing window. Note that all other windows will update at the same time.

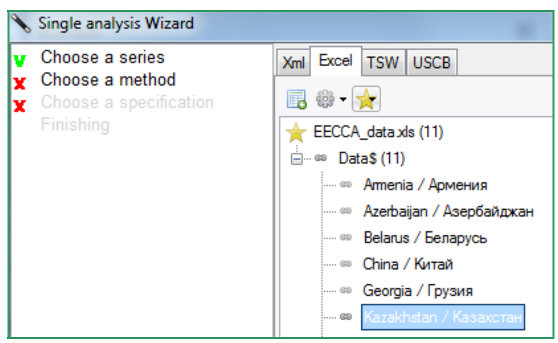
There you may initiate a new single processing by selecting **Seasonal Adjustment/Single Analysis/New**. If the *Workspace* panel already includes an active specification, the new single processing will correspond to that specification; otherwise a small dialog box will invite the user to select the specification.

Image 48 summarizes different steps and alternatives to follow in order to create a single analysis, with a user-defined specification, for example with TRAMO/SEATS.

The fourth option for single processing is to use the *Wizard* by means of the Main menu item (image 49): **Seasonal adjustment/Single Analysis/Wizard**. In the instructions for selecting regression variables, in the *Wizard*, you choose the series for seasonal adjustment and either create your own specifications or use the default specifications. When you finalise the *Wizard*, SEATS will automatically adjust the series.

The industrial production index of Kazakhstan plays the role of an example time series in the following instructions for single processing. In table 13 you can see the applied user-defined specifications. This example is an illustration and doesn't suggest an optimal treatment of this particular time series.

Image 49
A screenshot from the single processing Wizard



This example makes use of the TRAMO/SEATS method with a user-defined specification defined via the menu **Seasonal Adjustment/Specifications**. Demetra+ calls the single processing "TramoSeatsSpecDlg" to identify the processing task.

Table 13
User-defined specification used in the example of single processing

Options	Set	Definition
Transformation→ Function	Auto	Program will test for log/level specification
Calendar effects→ Trading days →Type	Predefined	
Details→ Trading days	td2	Working-day and leap year
Pretest	True	Program tests whether the effect is significant or not.
Easter effect → option Duration.	Pretest 6	Program tests whether the effect is significant or not. Default duration of Easter
Arima modelling → Automatic modelling → Enabled	True	The program automatically identifies the orders of ARIMA model
Outlier detection → Enabled	True	The program automatically detects the outliers

By clicking ok, **a new specification appears** in the *Workspace* panel, under the **Specifications/ TramoSeats** tree. By default, Demetra+ calls it "TramoSeatsSpec-1". To **activate the specification right-click** on it. Then, **double click on the series** of Kazakhstan in the *Browsers* panel to launch the adjustment.

Multi-processing

Select approach

Demetra+ can perform seasonal adjustment of a potentially large set of time series. This kind of a task is called "multi-processing".

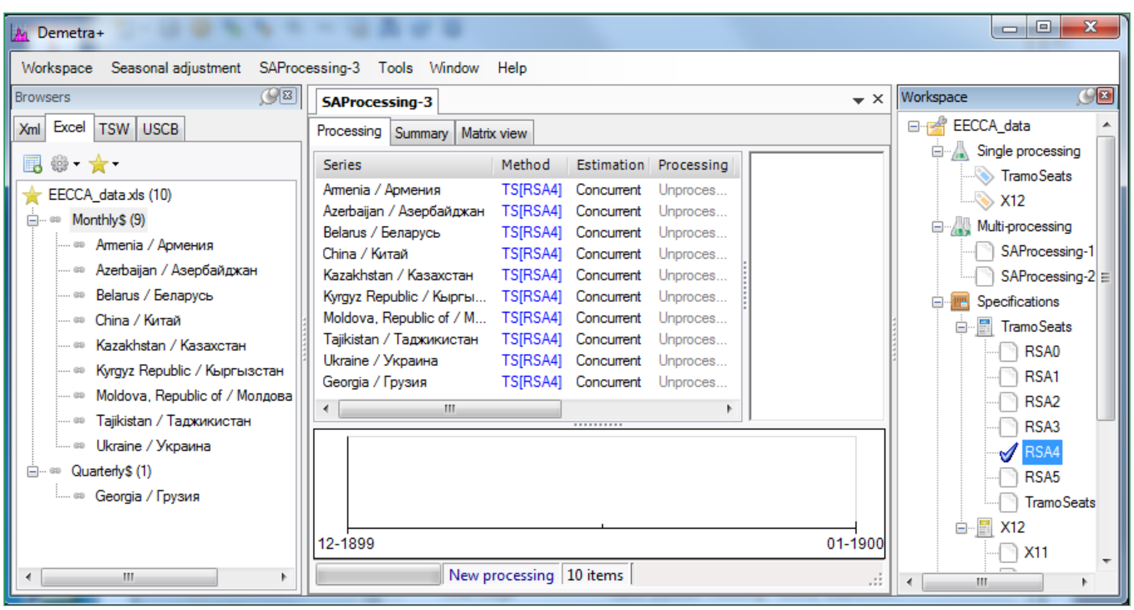
In multi-processing Demetra+ can adjust a set of time series with varying specifications. To start multi-processing, first define the methods and specifications to be applied for the whole set of series. The software provides some different ways for doing so.

You may either apply the currently active specification you've been using or you may use the *Wizard* which allows you to define the specifications for the series step-by-step.

Select specifications and regression variables

In Demetra+, TRAMO and Reg-ARIMA will apply the regression variables based on the selection of the specification. To define specifications for multi-processing, first **activate either a pre-defined or a user-defined specification** from the *Workspace* panel on the right side. Create a new multi-processing template by selecting **Multi-processing/New** from the Main menu. Select the series by **dragging and dropping them** into the multi-processing window. Demetra+ will apply the chosen active specification for adjusting these series.

Image 50
Creation of multi-processing with a pre-defined specification



The example of multiprocessing makes use of industrial production indices of the countries of Eastern Europe, Caucasus and Central Asia as well as China. Similarly to the example of single processing, this part purely aims to give instructions for the process and doesn't imply any optimal choices for adjusting these series.

First, the industrial production indices of these countries are adjusted with the specification RSA4 of TRAMO/SEATS. It includes a test for the working day effect, outliers and Easter, automatic model identification and pre-adjustment for leap year. First **activate the RSA4 specification** in the *Workspace* panel. Then **move the series into the multi-processing window** by dragging and dropping the heading for the set of series or their individual names from the *Browsers* panel (image 50).

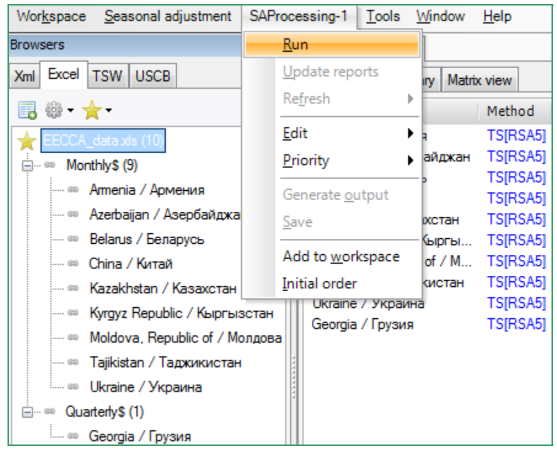
Launch seasonal adjustment

To launch the processing using the defined specification, choose **SAProcessing-n/Run** from the Main menu (image 51).

The other option to launching multi-processing is to use the *Wizard* (image 52). The *Wizard* guides the user in selecting specifications. It also offers the possibility to use specifications otherwise not part of the *Workspace*.

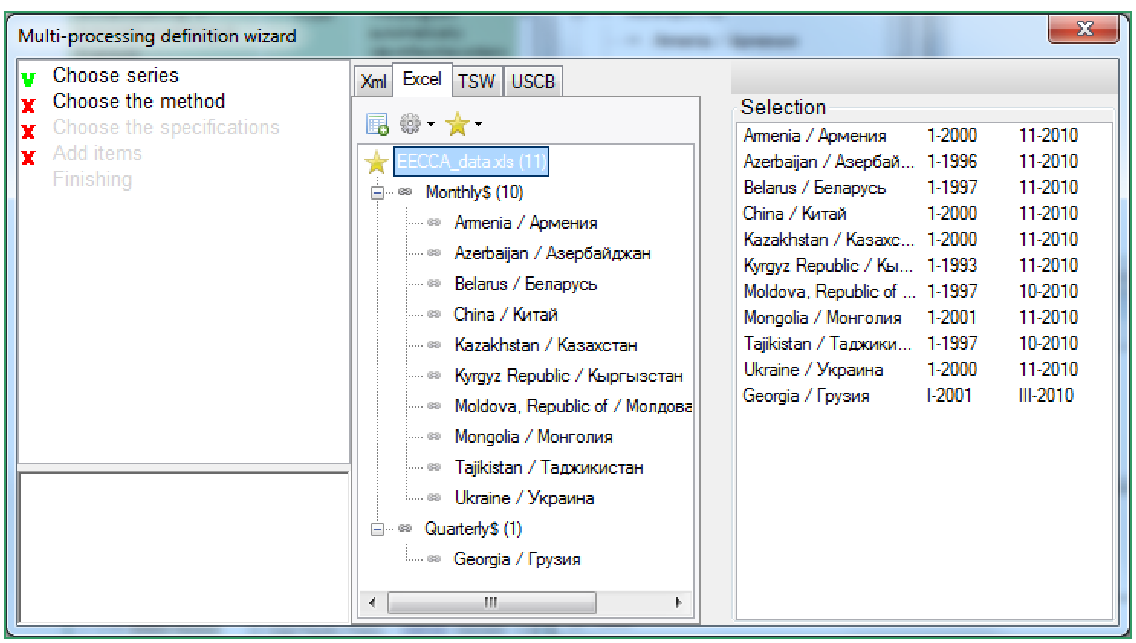
Follow through the steps of the *Wizard* and select your specifications for seasonal adjustment. The process is similar to that of single processing. At the last stage of the *Wizard*, Finishing, you may modify the name of the multi-processing, by default called the SAProcessing-n, where "n" refers to the number of processes. You can also add the multi-processing to the current *Workspace* for future re-use.

Image 51
Launching multi-processing



In the last part of the *Wizard*, you may decide to automatically start the execution when the *Wizard* is closed. You can go back to the first step of the *Wizard* at any time if you need to add other series with other specifications. You may find further details in the *Demetra+ User Manual* which explains the different functions. In chapter 4 we'll be discussing in more detail how to analyse and refine the results of seasonal adjustment.

Image 52
Using the Wizard to create multi-processing



CHAPTER 4

Analysis of the results

Introduction

In this chapter we'll describe the third phase of seasonal adjustment which includes analyses of the quality of seasonal adjustment and fine-tuning of the results. We'll introduce and explain the main quality diagnostics displayed by Demetra+, in particular for TRAMO/SEATS. Our aim is to help open the “black box” of the process of seasonal adjustment performed by the software, in order to support achieving robust results of seasonal adjustment. We'll be paying attention to the following quality issues:

- Appropriateness of the identified model and components.
- Number and type of outliers.
- Stability of the seasonal component.
- Absence of residual seasonality and residual calendar effects.
- Magnitude of the possible phase delay.

Seasonality is not a solid and precise fact but is identified based on hypotheses about the underlying conditions and models. The purpose of quality diagnostics is to reveal any essential weaknesses in the results of seasonal adjustment in order to prevent the use of misleading results that could lead to false signals about the economy. The automatic procedure of TRAMO/SEATS and X-12-ARIMA is quite reliable and, thus, useful for adjusting a large number of series. Especially with a limited number of important series it's of utmost importance to read the quality diagnostics well and with thorough consideration.

The output of Demetra+ presents the statistical and mathematical properties of the identified model and components. It includes a wide range of quality diagnostics that reflect the different approaches of TRAMO-SEATS and X-12-ARIMA. However, they have also a number of common quality diagnostics. In addition to numeric diagnostic tests, Demetra+ provides the user with a variety of readily available illustrative charts.

Smoothness of the seasonally adjusted series is not a quality measure. On the contrary, the irregular component is a part of the seasonally adjusted series. Ideally the number of outliers would be relatively small, and they shouldn’t be unevenly distributed in the series. Modelling problems are more likely with a short time series which includes many outliers. Longer time series would be particularly helpful for identifying the seasonal pattern from a highly volatile time series. Nevertheless, high volatility is a natural attribute of many time series of the emerging economies or quickly evolving industries.

Careful assessment of the seasonally adjusted data includes analysis of the stability of the seasonal component. Demetra+ reports the results of several quality diagnostics designed for this purpose. These include statistical tests and graphical diagnostics that depend on the chosen seasonal adjustment method. The M-statistics of X-12-ARIMA are explained in the *Demetra+ User Manual* and for instructions on the interpretation of the SEATS results see Maravall and Pérez (2011).

The residuals of the Reg-ARIMA offer a useful tool for verifying that the seasonal adjustment is adequate. In theory, the residuals should be random and not include any seasonality. On the other hand, Demetra+ also provides a test for confirming that there’s no residual seasonality in the seasonally adjusted data either. Seasonal adjustment should remove the seasonal component completely.

Phase delay describes the possible change of timing in the series resulting from seasonal adjustment. The possibly changed timing of turning points due to seasonal adjustment is important to recognize and communicate to the users of statistics.

The overall quality indicators displayed by Demetra+ help draw attention to the most problematic series. Where the quality diagnostics question the validity of the results, or indicate possible problems, modify the specifications and readjust the series.

Sometimes, the quality diagnostics indicate features which are problematic for a standard seasonal adjustment process. For example, some highly non-linear series don’t allow identification of a model with acceptable diagnostics, not even by shortening the series. A dominant irregular component or a large number of outliers could hide the seasonal model. In addition, inconsistent adjustments of overlapping time spans would alert about severely unstable seasonality. According to *the ESS Guidelines on Seasonal Adjustment*, these cases require consulting the literature, manuals and experts in order to develop an adequate solution (Eurostat, 2009).

Single processing

Main Results

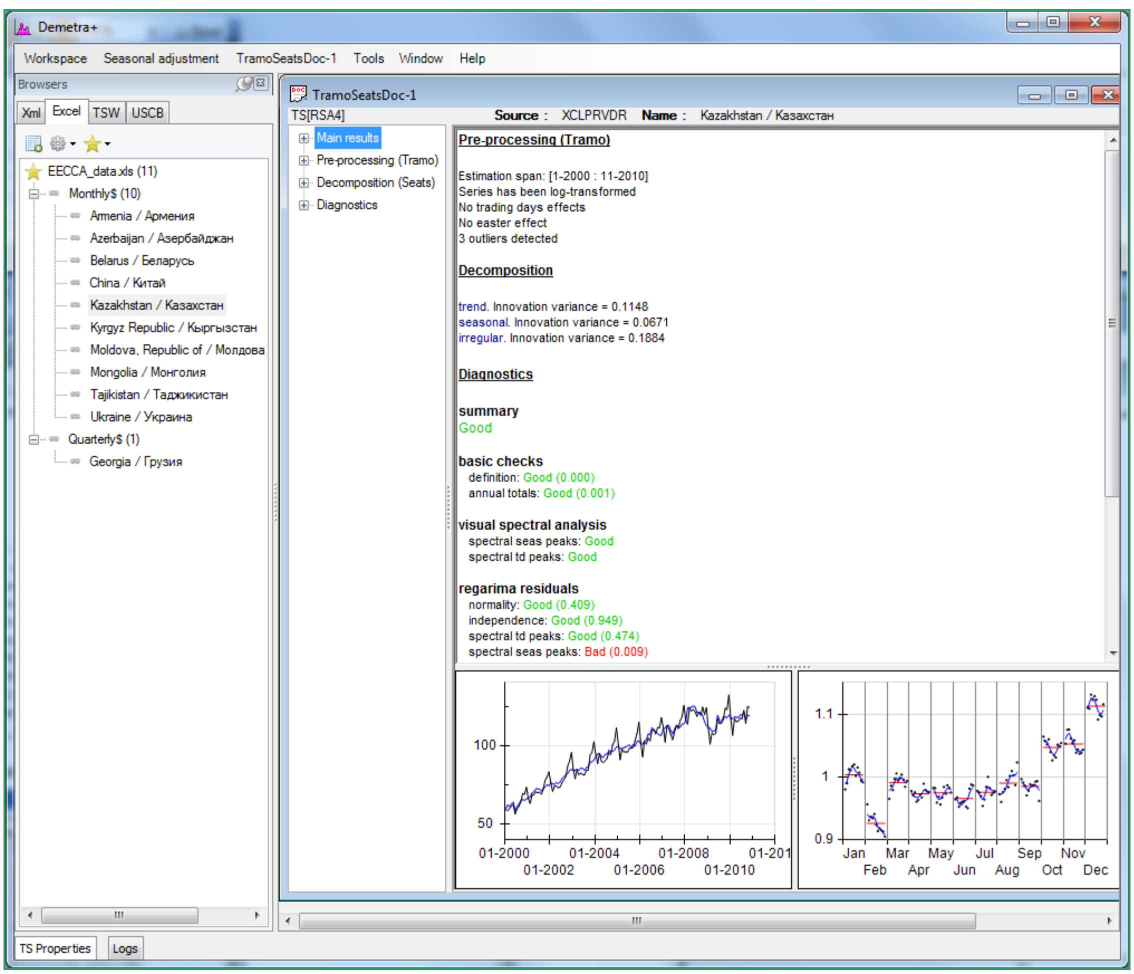
Demetra+ provides comprehensive and detailed tests for analysing the quality of seasonal adjustment. The available tests vary between the chosen adjustment methods. Demetra+ presents all the results of seasonal adjustment in four main parts: Main results, Pre-processing, Decomposition and Diagnostics. For users of X-12-ARIMA, the software displays the results in a similar way expressed with the terminology of X-12-ARIMA where pre-processing is replaced by the term Reg-ARIMA and decomposition by X-11.

The Main results contains a short description of the model used and of the quality of the seasonal adjustment. Image 53 displays the results obtained in seasonal adjustment of the industrial production index of Kazakhstan with the specification defined in chapter 3.

The Main results includes the following information about the adjustment process:

- The estimation span used for identifying the seasonal pattern was from January 2000 to November 2010.
- TRAMO/SEATS applied log-transformation to the series so that the data would more closely meet the statistical assumptions and to help fit the model. With an additive series the model is directly fitted, but the multiplicative series is first log-transformed to turn it into an additive form for the time of fitting the model.

Image 53
The view of results obtained by single analysis



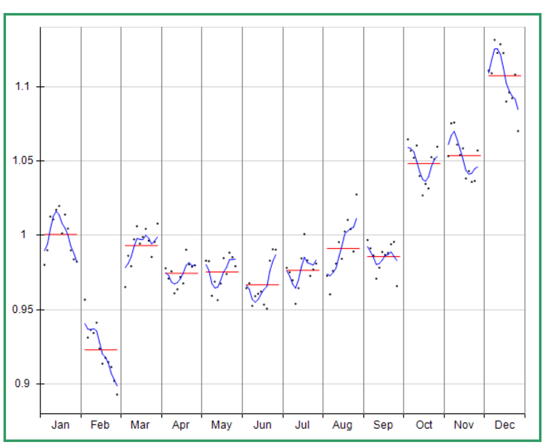
- Neither working day nor Easter effects were significant. Leap year effect was not present.
- The outlier detection procedure detected three outliers in the series.
- The variance of the seasonal and trend-cycle component innovations was lower than that of the irregular component meaning that the assumption of canonical decomposition holds.
- The quality of seasonal adjustment satisfied all diagnostics, except for the visual test on the spectral seasonal peaks. This warns about a possible misspecification of seasonality; an issue we'll be addressing later in this chapter.

If you wish to limit the amount of results displayed, you can do so by unselecting items for either TramoResults or X12Results at **Tools/Options/ Default SA processing output**.

Visual check

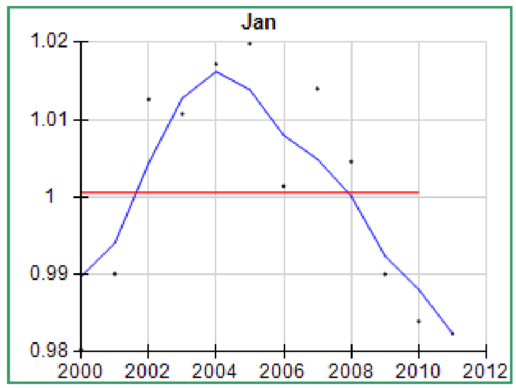
Demetra+ displays several charts and tables under the Main results. On this page the S-I ratio chart provides one of the most useful descriptive views of decomposition (image 54). The S-I chart illustrates unstable or moving seasonal factors, if any. It's useful for identifying seasonal breaks and periods with high variability. Seasonal breaks may cause problems with moving averages and identifying trading day and Easter effects as well as for fitting an ARIMA model. They are problematic to treat: You could prepare user-defined variables or adjust separately the two parts of the series. The latter would require several years of data before and after the break and is an option only for treating a historical break.

Image 54
S-I ratio chart



By double-clicking on a specific period, Demetra+ opens a more detailed view of the seasonal pattern of one single month or quarter. The high degree of changes in the observations for a month reflects the presence of moving seasonality (image 55), and abrupt changes would indicate seasonal breaks. For example, in the beginning of the time series, the values of seasonally adjusted data for February were lower than the raw data, and vice versa towards the end of the series. However, rather high volatility of data would be natural to emerging areas of the economy.

Image 55
A focused S-I ratio chart for January



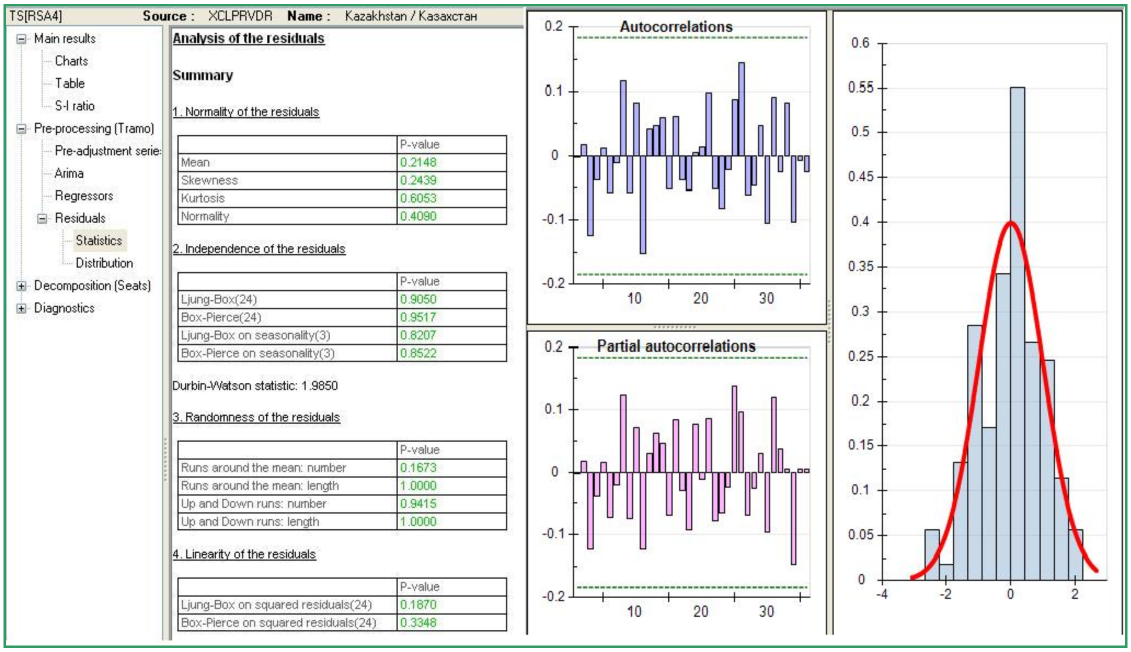
The Pre-processing displays the properties of the pre-adjustment step and contains the information listed here:

- Statistical properties of the ARIMA model used in seasonal adjustment.
- Regression variables and their coefficients.
- Tables of the pre-adjusted series.
- Residuals, with a complete analysis of their statistical properties, based on the diagnostics produced by TRAMO.

The residuals of the model are useful for checking the quality of seasonal adjustment. The residuals should be independent and random, and thus follow the normal distribution. The average should be zero; they shouldn't contain information, such as seasonality and they should be distributed according to a normal distribution.

Different ARIMA models could fit the same series. Many statisticians recommend choosing the simplest model with the smallest number of parameters but with a satisfactory fit. For the series of Kazakhstan, TRAMO has selected an Airline model (0,1,1)(0,1,1). In general, the automatic model identification should produce

Image 56
Test statistics and distribution graph of the residuals



satisfactory results, and you shouldn't need to manually define the model. The results may require refining and changes in the specification.

The diagnostic tests on residuals don't suggest any statistical problems (image 56). All null hypothesises of statistical tests are accepted at five per cent significance level. Hence, the distribution of residuals is random, normal and independent, which means that they don't include any information, and they fulfil the theoretical requirements.

Significance level is the criterion used for testing an assumption, in other words the null hypothesis about the data. An experienced user may change the significance levels applied by Demetra+ by **Tools/Options/Diagnostics**.

The outlier detection has identified one transitory change in June 2000 and two level shifts in June 2009 and February 2008. A large number of outliers relative to the length of the series could result in over-specification of the regression model. Re-modelling of the series by reducing the number of outliers could lead to better results. Changing the identified outliers requires careful consideration.

Collected information to verify the causes of outliers is especially important at the end of the series where the type of outliers is uncertain, as the next observations are not yet available. The change of the type of outlier later may lead to large revisions.

Models applied

The Decomposition contains information about the ARIMA models applied by SEATS. The part presents, for SEATS, the properties of the Wiener-Kolmogorov filters which are used to extract the components from the original series. Decomposition requires that the components are uncorrelated. If you are using X-12-ARIMA, Demetra+ provides the detailed tables namely, A, B, C, D and E which contain the results of the consecutive stages of the X-11 procedures of decomposition as applied by X-12-ARIMA.

SEATS identified an ARIMA model for each component of the series of Kazakhstan which are the trend-cycle, seasonal and irregular component. Demetra+ presents the mathematical formulas for each ARIMA model of a component.

In this case, the models of the components obtained by SEATS are:

$$\textit{Main model}: \Delta\Delta^{12}x_t = (1-0.12B)(1-0.55B^{12})a_t, \quad \sigma_a^2 = 1$$

$$\textit{Trend model}: \Delta^2 TC_t = (1+0.049B-0.95B^2)a_{TC,t}, \sigma^2_{TC,t} = 0.115$$

$$\textit{Seasonal model}:$$
$$(1+B+B^2+...+B^{11})s_t = \begin{pmatrix} 1+1.72B+1.99B^2+2.04B^3 \\ +1.88B^4+1.61B^5+1.29B^6 \\ +0.95B^7+0.64B^8+0.33B^9 \\ +0.14B^{10}-0.19B^{11} \end{pmatrix} a_{s,t},$$
$$\sigma_s^2 = 0.07$$

$$\textit{Irregular}: \quad \textit{White noise } (0\,;0{,}19)$$

Where, Δ is the difference operator, B is the lag operator ($B1x_t$ = x_t-1), σ^2 is the proportional variance of innovations of the component.

Demetra+ presents the resulting components under the Stochastic series which includes the seasonally adjusted series as well as the trend-cycle, seasonal and irregular components. The latter contains the transitory component if it's identified.

Here, under Model-based tests, Demetra+ also reports the **cross-correlations** between the components of the series. The table of cross-correlations provides the user with useful information for testing whether the estimators or actual estimates of components correlate with each other or not. The results of cross-correlation are a way to test the assumption of "orthogonal" (uncorrelated) components. Table 14 presents a cross-correlation table produced by SEATS for the industrial production index of Kazakhstan. In Demetra+ the colour of the p-values signals the result, e.g. a green p-value indicates insignificant cross-correlation.

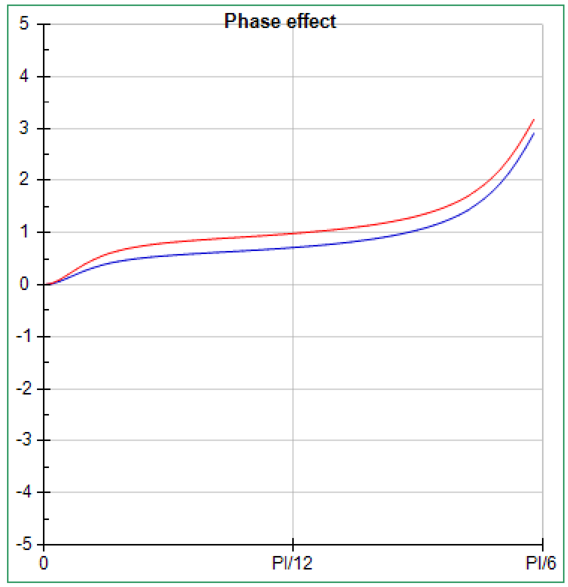
Table 14
Cross-correlation of results

	Estimator	Estimate	P-Value
Trend/Seasonal	-0.1250	-0.1504	0.8018
Trend/Irregular	-0.0450	-0.0856	0.7311
Seasonal/Irregular	0.0446	0.0195	0.5900

Analysis of economic development pays a lot of attention to accurate and timely detection of turning points. For this purpose, Demetra+ displays some advanced tools, such as the estimation of **phase delays** or phase shifts. A phase delay is the time shift between the adjusted (seasonally adjusted and trend-cycle) and the unadjusted series. This means that the timing of turning points has changed so that they occur either earlier or later in the adjusted series than in the original. A positive phase delay means that the adjusted series has shifted backwards, and a negative phase shift that the adjusted series has shifted forward in time.

For the series of Kazakhstan, the chart shows that, at the end of the series, the decomposition generates a phase delay effect of less than one month on the seasonally adjusted and the trend-cycle series (image 57).

Image 57
Phase-delay effect of seasonal adjustment



Quality diagnostics

The Diagnostics presents detailed information on the seasonal adjustment procedure. This information, often purely descriptive, Demetra+ computes the same way for TRAMO-SEATS and for X-12-ARIMA. Demetra+ divides the diagnostics into five main parts: presence of seasonality, spectral graphics, revision history, sliding spans and model stability analysis.

The first part of diagnostics offers the results of tests on the **presence of seasonality** in the original series and in the residuals. The series has to contain a seasonal pattern; otherwise there's no need to perform seasonal adjustment at all.

Demetra+ provides the user with several **seasonality tests**, such as Friedman and Kruskall-Wallis, a test for stable seasonality and a moving seasonality test. Information about residual seasonality is available at the end of the Diagnostics.

Image 58
Presence of seasonality in the series of Kazakhstan

Evolutive seasonality test

	Sum of squares	Degrees of freedom	Mean square
Between years	0.0028	9	0.0003
Error	0.0117	99	0.0001

Value: 2.6270
Distribution: F-stat with 9 degrees of freedom in the numerator and 99 degrees of freedom in the denominator
P-Value: 0.0091
Moving seasonality present at the 5 per cent level

Combined seasonality test

Identifiable seasonality present

Residual seasonality test

No evidence of residual seasonality in the entire series at the 10 per cent level: F=0.1024
No evidence of residual seasonality in the last 3 years at the 10 per cent level: F=0.4541

For the index of industrial production of Kazakhstan, image 58 shows the results of seasonality tests. According to them, seasonality is clearly present in the original series which fulfils the main prerequisite for seasonal adjustment. The fact that there's no residual seasonality in the entire duration of the series, nor during the last three years of the series, reflects good conditions for identifying the seasonal pattern.

However, some **moving seasonality** is present in the series of Kazakhstan, at a five per cent significance level. The presence of moving seasonality is not a surprise, considering the S-I ratio chart discussed before. This test confirms the findings of the visual analysis.

The second part of the Diagnostics, presents **spectral graphics**, namely a Periodogram and an Auto-regressive spectrum. These spectral graphics analyse the residuals, irregular component and seasonally adjusted series for remaining seasonal or trading day effects. They check for the presence of peaks at the seasonal frequencies (the grey lines

Image 59
Spectral graphics of residuals obtained by TRAMO

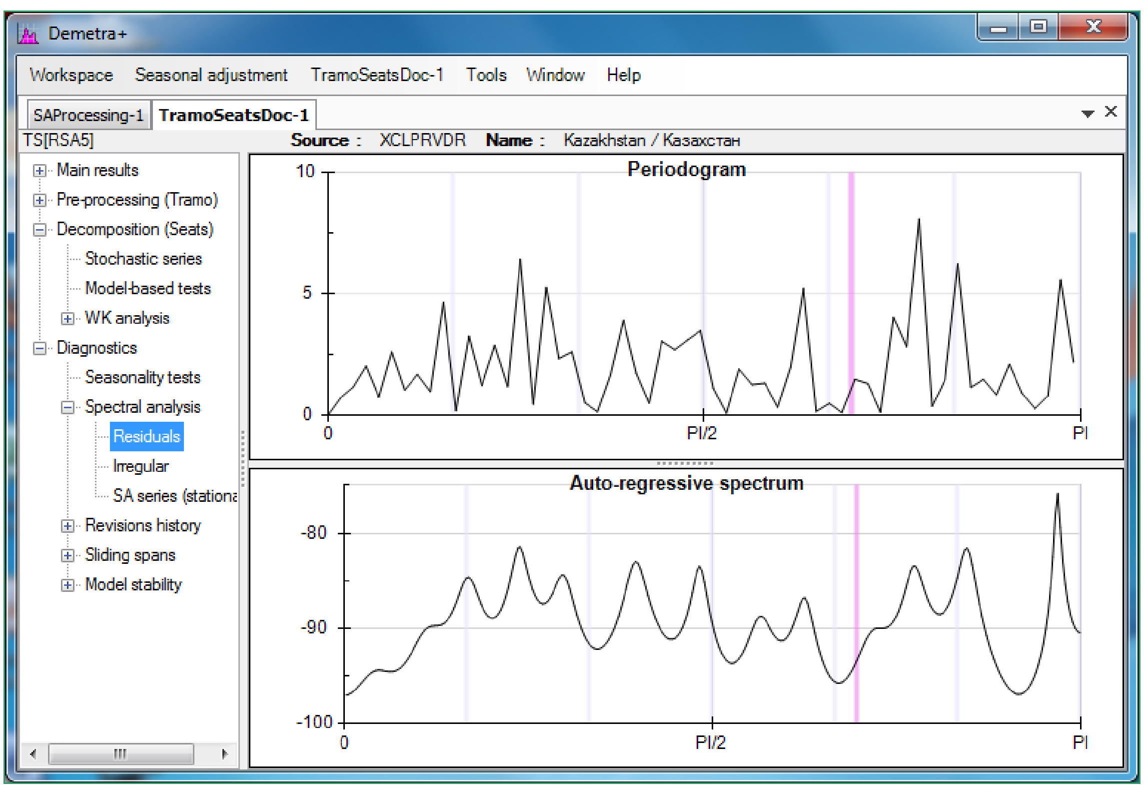
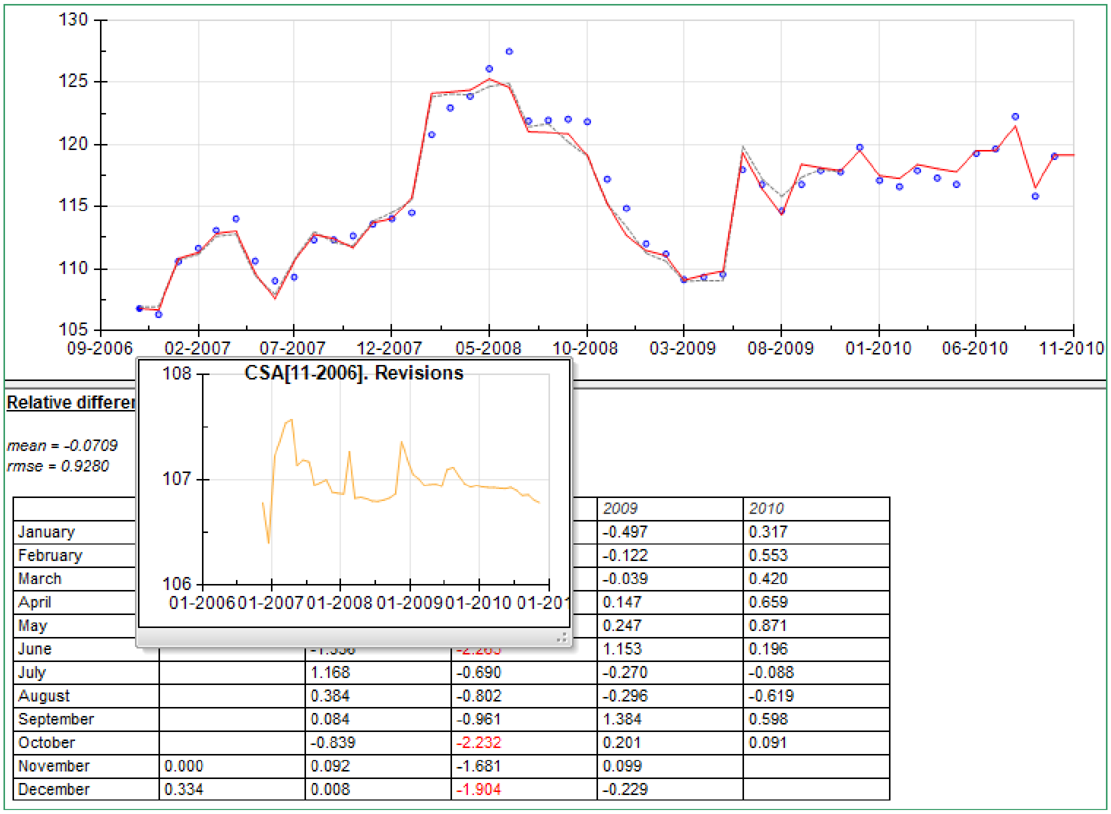


Image 60
Revision analysis



in Demetra+) and at the trading day frequency (the purple line in Demetra+). Peaks at the seasonal frequencies of an adjusted series mean that the filters used in the decomposition aren't well adapted to the series or not to a large part of it. Peaks at the trading day frequencies could indicate that the regression variables of the model don't suite well the series or that the calendar effects change too much to be captured by the fixed regression effects applied for the whole duration of the series. If remaining seasonality is present reconsider the model specification, regression variables or the time span used for modelling.

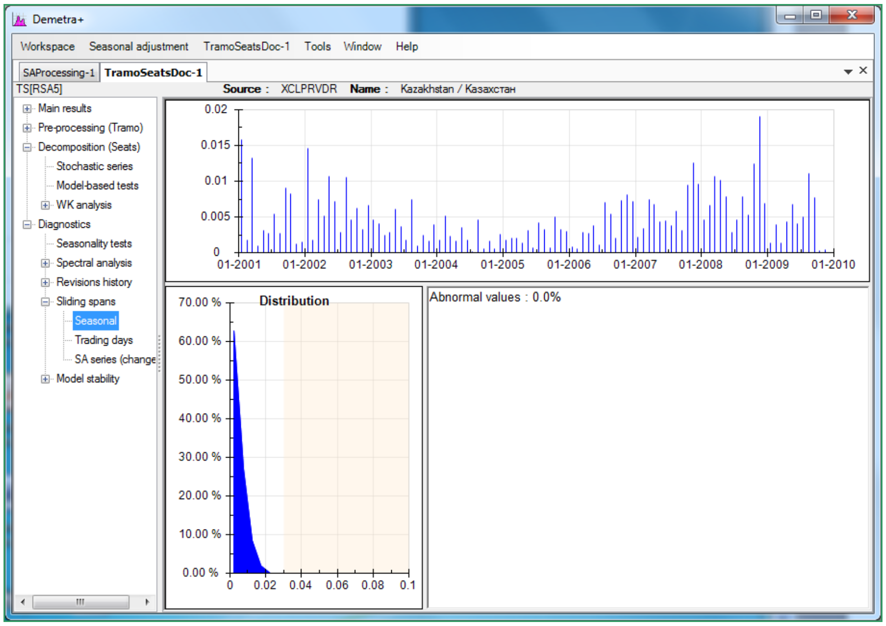
The series of Kazakhstan doesn't show any indication of **residual seasonality** or residual calendar effects. In other words, image 59 depicts no spectral peaks at the seasonal or trading day frequencies. Nevertheless, a peak is clearly visible in a very short-term frequency, but this isn't easy to interpret and is unimportant in terms of seasonal adjustment.

The third part of the Diagnostics, presents the **revision history** of the series. It analyses what kind of revisions adding new observations at the end of the series causes. It presents charts both for the seasonally adjusted and trend-cycle series. Demetra+ displays the revision history for both methods, TRAMO/SEATS and X-12-ARIMA.

Each blue circle on the chart in Demetra+ depicts the initial adjustment when this point is the last observation. The red curve presents the final results. (Image 60.) The analysis starts by estimating the model for the complete time span. There on, the time span is shortened progressively and the decomposition re-estimated. For each period, a series of successive estimations is obtained. By default, it re-estimates only the parameters of the model. However, the program allows a complete re-estimation and a re-identification of the outliers if the option is changed.

By clicking on a dot in the graph, for example the observation of November 2006, a pop-up window appears. It shows the successive estimations for November 2006 for the different time spans [t0, 2006-11], [t0, 2006-12]...[t0, 2010-1].

Image 61
Sliding spans analysis



The pop-up window for the industrial production index of Kazakhstan confirms that revisions are not significant after three years. In the example series, the sudden change in February 2008 is linked to the level shift that appeared at the same time.

This part also contains a revision history table that presents the differences between the first estimates and the last estimates for the last four years. If the decomposition is additive Demetra+ displays absolute revisions; otherwise it uses relative differences. To enable quicker analysis, Demetra+ displays the largest differences in red. Any red observations are, in absolute terms, larger than two times the root mean squared error of the absolute or relative revisions. The series of Kazakhstan includes four observations in 2008 which exceed the given critical limit. Three of these abnormal values are, in fact, outliers.

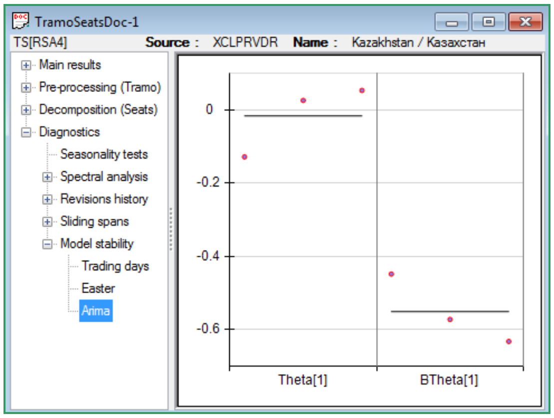
The fourth part of the Diagnostics, includes the **sliding spans** analysis, originally used in the X-12-ARIMA. It is particularly useful for a series with changes in seasonality or a large number of outliers. From the analysed time series, the program extracts time spans with the length of eight years. In this example it extracts four time spans which are separated from each other by one year. When an additive decomposition is used the sliding spans analysis is based on absolute differences. The threshold to detect abnormal values is set to three per cent. Any larger value is unstable. As seen in image 61, according to the sliding spans analysis of the series of Kazakhstan, the seasonal factors appear to be stable since none of the relative differences exceeds three per cent. The sliding spans are computed for the seasonal component, the trading day effect, if any, and the seasonally adjusted series.

The results are stable if adding or removing observations doesn't cause a lot of change in the results. Stability is generally a good indicator of the quality of adjustment. However, for some time series, even the best possible seasonal adjustment may be unstable. In such a case, balancing between the quality and the extent of revisions will be a challenge. You could try changing the model specification if you detect many unstable estimates. In chapter 5 we'll discuss revision policies in more detail.

The last part of the Diagnostics provides **model stability** analysis which calculates ARIMA parameters and coefficients of the regression variables for different periods. It shows the results in a visual form. The model stability analysis computes the results on a moving window of eight years which slides by one year at a time. The displayed points correspond to the successive estimations.

The model stability analysis for Kazakhstan includes the time spans for 2000-2008, 2001-2009 and 2002-2010 (image 62). In the Airline model, the regular moving average parameter represents the structure of the trend-cycle component, whereas the seasonal moving average parameter reflects the seasonal component. The graphic shows an unstable regular moving average parameter, but a seasonal moving average parameter that evolves rather smoothly. The range of fluctuation for the two moving average parameters is of the same order. In the regular moving average parameter, the movement of values from negative to positive indicates a moving trend-cycle.

Image 62
Model stability analysis



The seasonally adjusted series is unobservable, which means that it needs to be estimated by applying estimators. Therefore, the estimation error is important, in particular for recent periods. SEATS provides approximate standard errors for the seasonally adjusted series and trend-cycle estimators. It provides further forecasts of these components jointly with their standard errors and calculates confidence intervals around the seasonally adjusted and trend-cycle series as well as their forecasts.[10]

The standard errors provided by SEATS reveal relevant properties of the seasonally adjusted series: When a large standard error characterizes the seasonal innovation, it indicates a highly moving seasonality; if it characterizes the estimation error, it would indicate imprecise estimation of seasonality; and if it characterizes the revisions, it would indicate unstable seasonality.

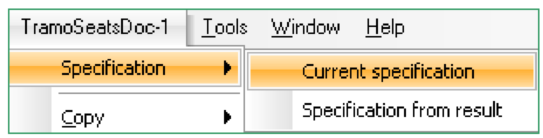
The *Demetra+ User Manual* includes comprehensive descriptions of the diagnostics for both X-12-ARIMA and TRAMO/SEATS.

Refine results

Performing seasonal adjustment is usually an exploratory process. The general approach is to start with a specification that gives more freedom to TRAMO and Reg-ARIMA so that they can try to find the best possible model. If need be, the user may progressively impose some constraints by modifying the specifications.

One of the main features of Demetra+ is that it allows an exploratory process of seasonal adjustment. For each round of the adjustment, the user may change any option in the specifications and will immediately see the effect on the results. It's very useful to test the options and their effects also to get to know the data and the possibilities offered by the software. Image 63 illustrates how to read the current specification used and to modify it by using the Main menu.

Image 63
Modification of results



After the following text on multi-processing, we'll be looking at how to do readjustment of the same time series later, and how to export the results to other software for further analysis, editing and publishing.

Multi-processing

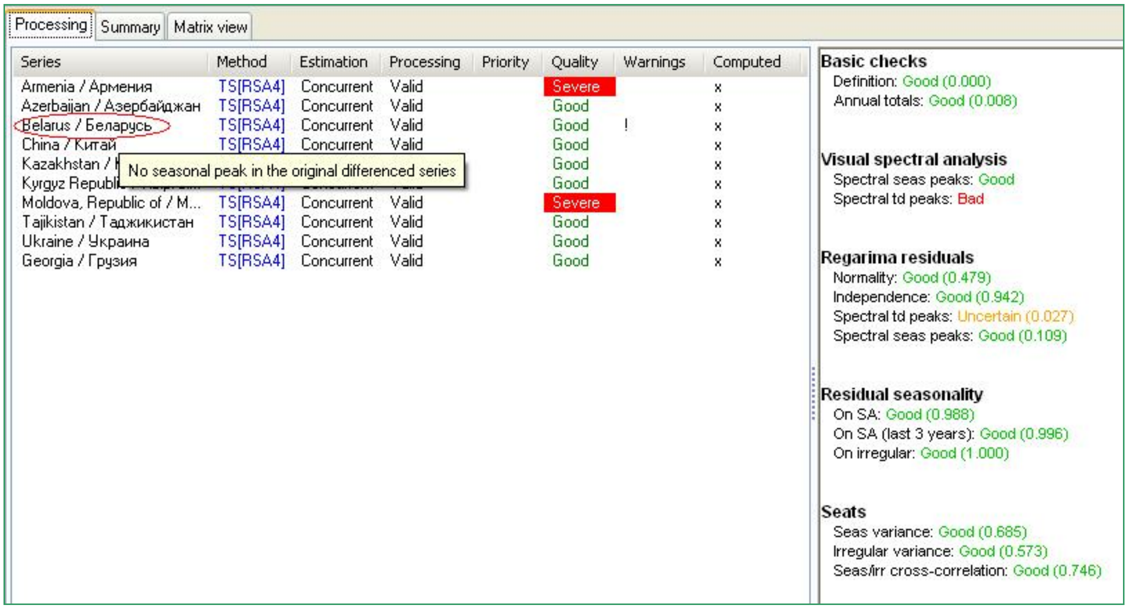
Main results

An overview of the main results presents the main conclusions of the multi-processing for each series (image 64). This window includes **an overall quality indicator**. You may refer to chapter 1 for more details about each level of quality (table 3).

This window may include some **warnings**, e.g. for a short time series, non-decomposable models (SEATS) or when the differenced series doesn't show any seasonal peaks. Demetra+ displays these warnings by adding a tip on the series in question. By a double click on the name of the series, you may open the quality diagnostics for that series.

10 For more information see: Moore, G.H., Box, G.E.P., Kaitz, H.B., Stephenson, J.A. and Zellner, A. (1981), *Seasonal Adjustment of the Monetary Aggregates: Report of the Committee of Experts on Seasonal Adjustment Techniques*, Washington, D.C.: Board of Governors of the Federal Reserve System.

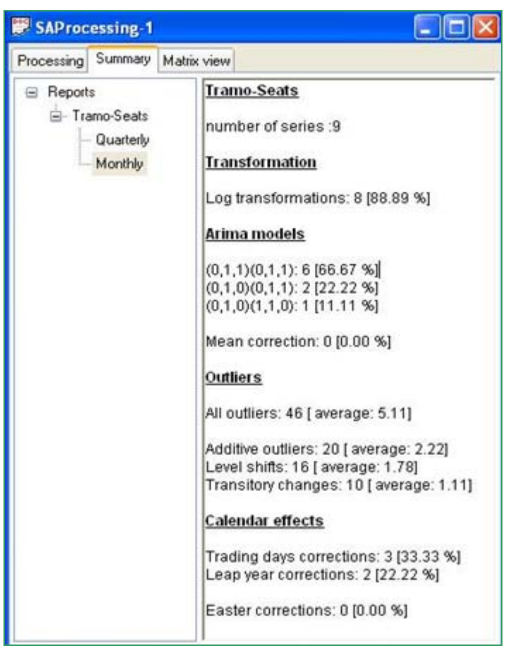
Image 64
Main results of the multi-processing



Models applied

The Summary gives general information on the **model used** for the set of series (image 65).

Image 65
Summary results for the set of series



TRAMO/SEATS has identified an Airline model for most countries (66.7%). TRAMO detected a trading day effect in 33.33%, and for nine monthly series, it did not detect any Easter effect. According to these results, calendar effects are not significant in most of these countries' industrial production indices. However, the series may still include calendar effects, i.e. the effect of specific national holidays. The number of outliers per series is around five. Given the volatile structure of some of the series used in the example, the number of outliers per series is at an expected level.

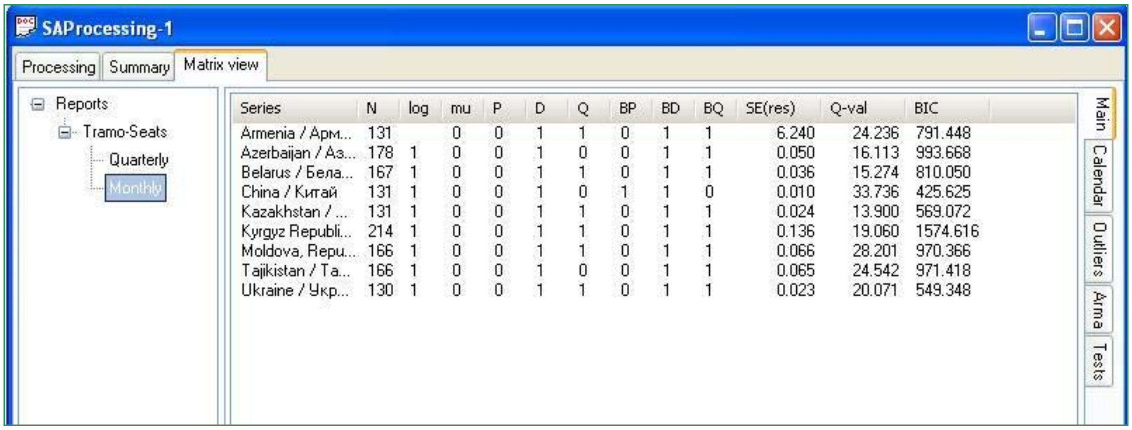
In image 66, the Matrix view contains **summary information** for each series about the main statistical properties of the ARIMA model used in pre-adjustment phase, the calendar specification results, the outlier structure of each series, the ARIMA parameter values and the significance tests. You may find here a table with the p-values of tests on residuals and other information, including annual discrepancies between raw and adjusted data as well as spectral visual peaks.

You can **copy** the matrices into the clipboard by the key combination **ctrl+c** and paste it with **ctrl+v** to another software. Note also that Demetra+ will display the details of a processing by a double click on the name of a series.

Quality diagnostics

Based on the first results obtained from the multi-processing (image 64), TRAMO/SEATS was able to perform seasonal adjustment well for all countries, with the exception of Armenia and the Republic of Moldova. For these two countries, the residuals obtained from the pre-adjustment step (TRAMO) may contain seasonal peaks since the corresponding diagnostics show severe overall

Image 66
Matrix view of Demetra+



quality. According to table 3, for statistical reasons the results are not acceptable.

According to a tip appearing for the industrial production index of Belarus, with specification RSA4, TRAMO/SEATS doesn't detect any seasonal peak in the original data. It could mean that there's no significant seasonality in the original series.

By **clicking on the name** of a series, **summary diagnostics** appear in the right part of the window, and the final seasonally adjusted series appears in the chart at the bottom of the window (image 66). By **double clicking on the name** of the series, Demetra+ offers **complete details** for the series.

When this window for the details is open, **changing the selected series** in the multi-processing window will automatically **refresh** the details according to the new selected series.

You may **sort** the multi-processing results by clicking on a column header. Sorting may be helpful when the processing contains many series.

Refine results

If it appears that the quality of a specific processing is insufficient (image 67), you should go to the detailed window and try to modify some options to get a better result.

Image 67
Information presented in the multi-processing window

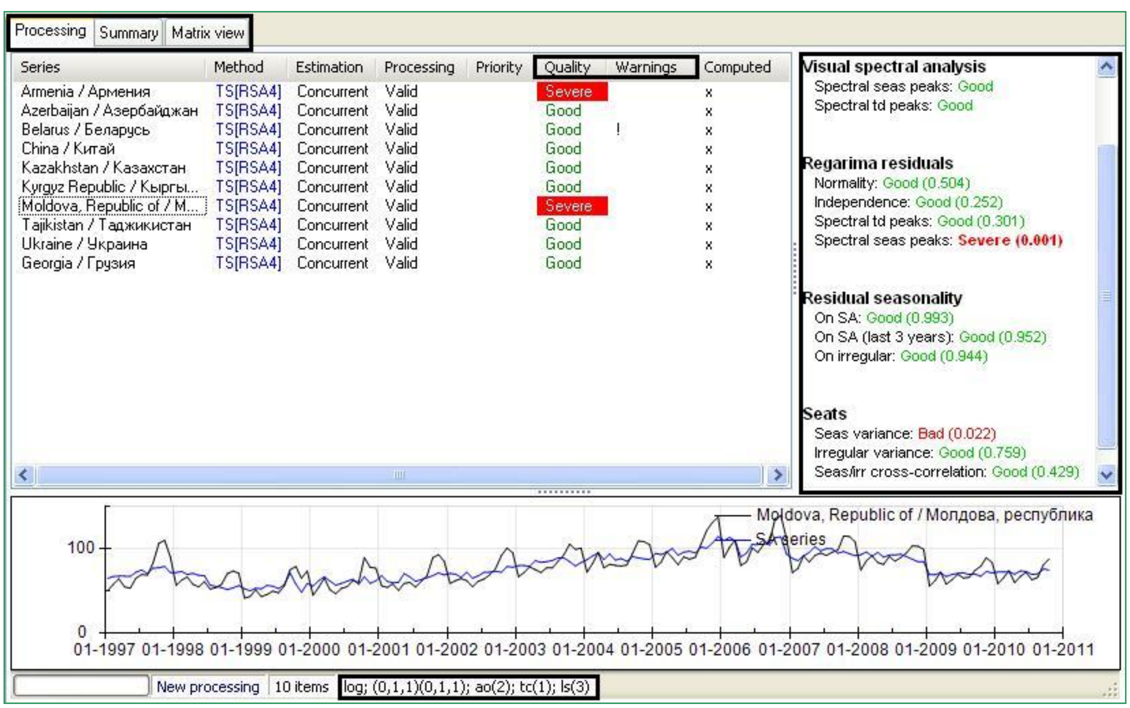
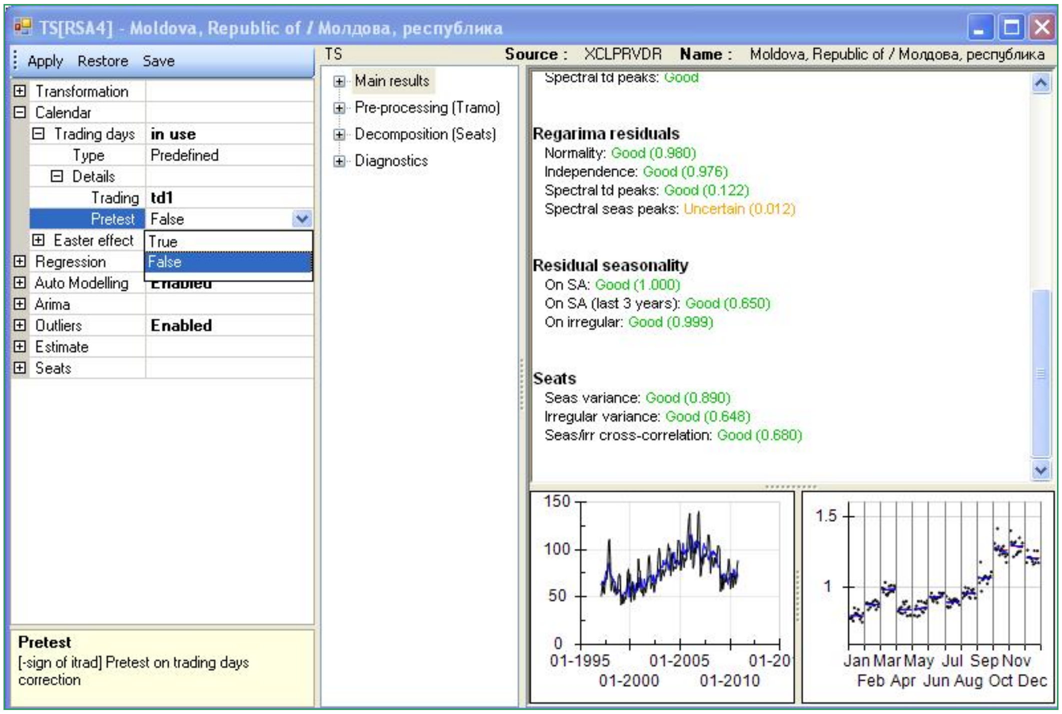


Image 68
Modification of the specifications in multi-processing



In the example shown for the Republic of Moldova, the chosen specification for trading day effects was not able to remove all the re-occurring effects.

The diagnostics indicated, among other things, the presence of spectral peaks on the seasonal frequencies and a high seasonal variance.

You may try to solve these types of problems by changing the trading day specification:

- Replace the trading day correction with working day correction (td1).
- Impose the pre-test by changing it to False.
- Apply a new specification.
- When the result is acceptable, save it to the multi-processing window by **clicking the Save button**.

You may also **add new series** into the multi-processing, by selecting **SAProcessing-n/Edit/Add** item from the Main menu.

Readjust results

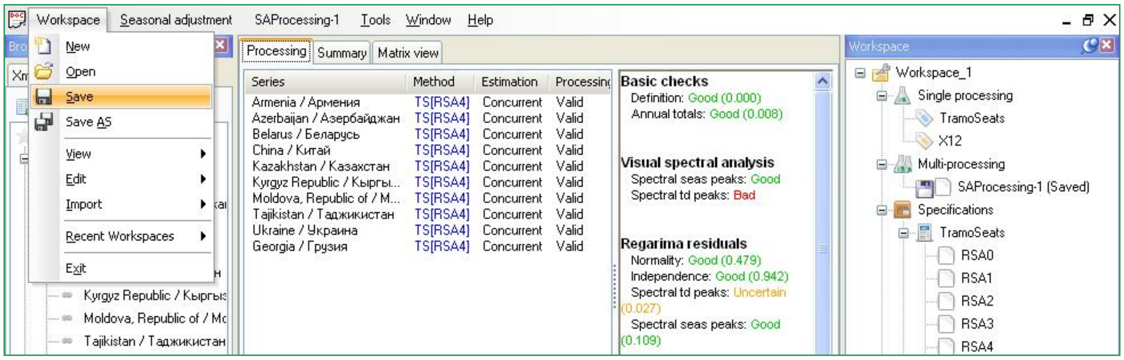
Usually, the need for performing seasonal adjustment repeats regularly depending on the frequency of data.

New observations and revisions refresh time series either monthly or quarterly. You may use the earlier processing as a basis for refreshing the seasonally adjusted series.

To be able to refresh a processing with new data later, follow the next rules:

- The series must come from the *Browser*s since Demetra+ wouldn't be able to refresh dynamically the series imported by dragging and dropping or by using the clipboard.
- Keep the identifiers of the data the same, including the path of the source file, names of the sheets and series.
- Add the earlier multi-processing to a *Workspace* and save it with the *Workspace*, in order to access it later. You may do this by selecting **SAProcessing-n/Add** to *Workspace* from the Main menu and then by selecting **Save** the *Workspace* (image 69).
- If the processing contains any user-defined regression variables or calendars, they need updating before readjustment, e.g. to add new observations.

Image 69
Saving the Workspace for future processing



When you get new observations and wish to refresh the earlier seasonally adjusted results, follow the steps given below (image 70):

- Update the source file with the new data in the folder where it's stored. Don't change the place of the data file in the computer.
- Open the previous *Workspace* in Demetra+.
- Double **click on the SAProcessing-n**, or other name given for it, under the Multi-processing node of the *Workspace*.
- Find the **Refresh** tool under the SAProcessing-n folder of the Main menu.
- Choose the refreshment strategy and generate output.

As we saw in chapter 1, in Demetra+ current adjustment (partial) refers to using fixed specifications with the exception of user-defined regression variables. In the concurrent adjustment, on the contrary, TRAMO/SEATS or X-12-ARIMA re-identifies the model, filters, outliers and regression parameters and the respective parameters and factors again. This corresponds to doing a new seasonal adjustment for the series.

Partial concurrent adjustment is a compromise between these two approaches. Demetra+ offers four different options for partial concurrent adjustment. If you select the option Parameters, it will fix the ARIMA model and the position of the outliers, but it will re-estimate all their parameters. The second option, Last outliers (+params) adds re-estimating last outliers to the previous option. The third option, All outliers (+params) adds re-estimating the position of all outliers to the previous approach. The fourth option, ARIMA and outliers (+params) additionally revises the ARIMA model. After selecting the refreshment strategy, TRAMO/SEATS or X-12-ARIMA performs the readjustment of the series immediately. Table 15 presents a summary of the refreshment strategies.

Export results

Demetra+ provides the user with several alternatives for exporting the results of seasonal adjustment for use in other software and in other formats. For multi-processing, it is able to export the results into different formats including *.txt, *.xls, ODBC or *.csv.

Image 70
Refreshment strategies

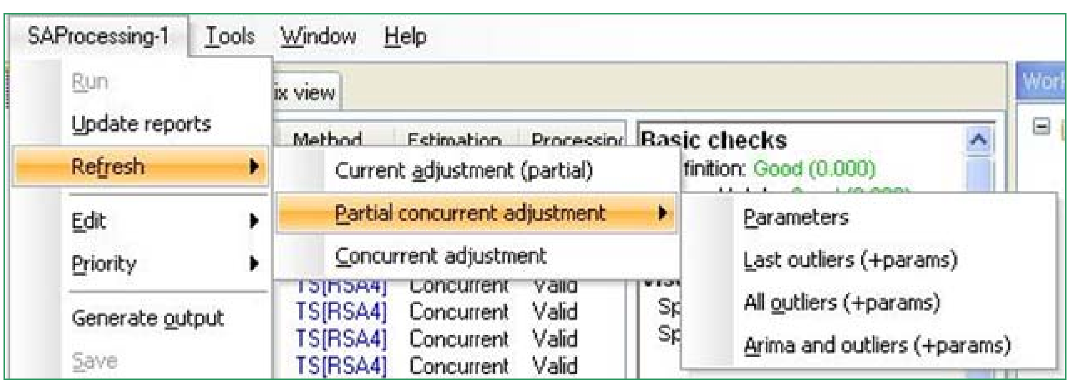


Table 15
Summary of refreshment strategies

		Outlier specification		**ARIMA specification** (p,d,q)(P,D,Q)		**Calendar Effects** (Trading and working day, and Easter Effect)
		Position	Parameters	Orders	Parameters	Parameters
Current adjustment (partial)		Fixed	Free	Fixed	Fixed	Fixed
Partial Concurrent Adjustment	Parameters	Fixed	Free	Fixed	Free	Free
	Last Outliers (+params)	Only last outlier free	Free	Fixed	Free	Free
	All outliers (+params)	Free	Free	Fixed	Free	Free
	Arima and Outliers (+params)	Free	Free	Free	Free	Free
Concurrent Adjustment		Free	Free	Free	Free	Free

To generate output, select **SAProcessing-n/Generate output** from the Main menu (image 71). This menu is visible provided that you've performed multi-processing. You may select the format of the output file and define some options for it (image 72).

Another alternative is to **copy and paste** the results from Demetra+. Most of the results displayed by Demetra+ are available for copying **by using the ctrl+c** and **ctrl+v** key combination or by **dragging and dropping** the data into other software.

After the processing, especially with smaller amounts of data, you may directly copy the results under **Main results/Table** in the *TS Properties* panel. By **right-clicking on the corner** of the table, a window opens and you can select **Edit/Copy all**. Then you can **paste** the results to the source file for further analysis.

Image 71
Generating output

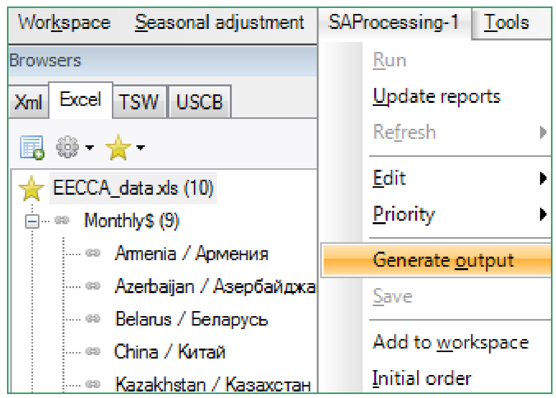
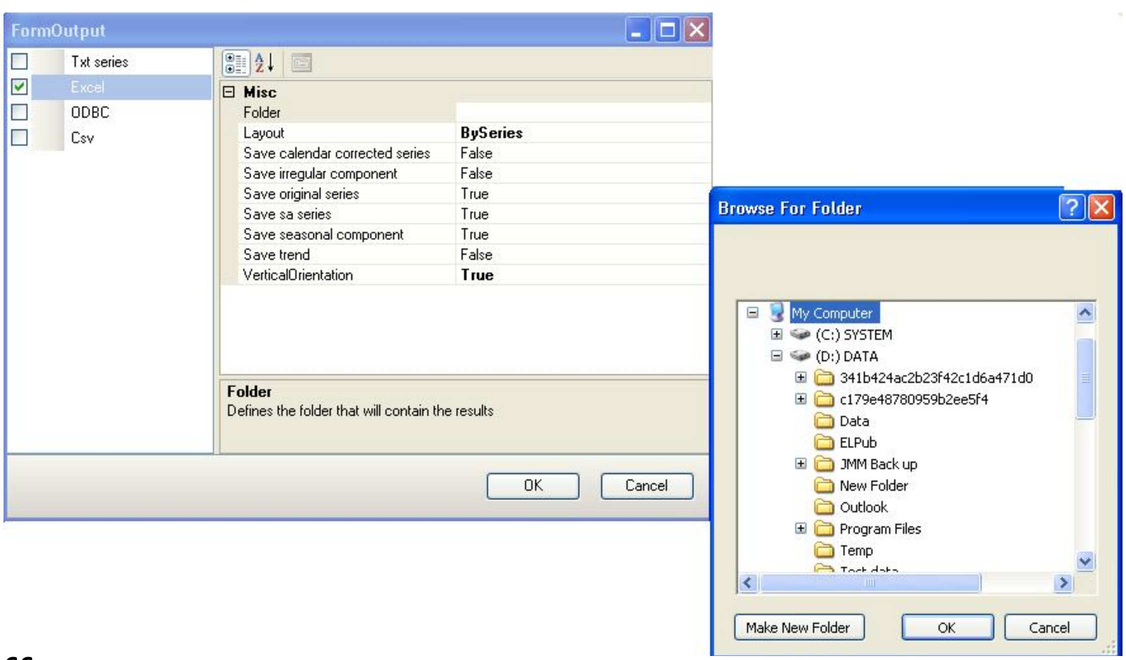


Image 72
Steps in generating output



CHAPTER 5

User communication

Introduction

Only used statistics are useful statistics - the principle applies to seasonally adjusted data as well. Seasonal adjustment process doesn't end when the quality of results has been confirmed. In this chapter we'll be looking at the fourth phase of seasonal adjustment dealing with how to support users of statistics in using seasonally adjusted data. We'll examine how to:

- Draft internal and external documentation.
- Define a seasonal adjustment policy.
- Revise seasonally adjusted data.
- Prepare data releases.
- Provide users with the necessary support.

Most of all seasonal adjustment aims to support the users of statistics in interpreting economic development. It removes the repeating patterns of data which may hide the underlying development. However, seasonal adjustment is a complex estimation technique that modifies the initial observations resulting in a theoretical understanding of reality.

Sufficient internal and external documentation on seasonal adjustment is particularly important because of the complexity of the method. With general explanations, the users can easily understand the technique and it helps interpret the data correctly. Without proper documentation, seasonally adjusted data could confuse the users. Internally, proper archiving of seasonally adjusted data and clear working instructions help produce robust results of seasonal adjustment.

A seasonal adjustment policy improves the quality and consistency of the national statistics. The policy makes decisions about a common software and the method to be applied. It takes a stance on the timing and methods of revising the seasonal models, treatment of outliers and contents of documentation. Therefore, an individual statistician doesn't have to make all these difficult decisions alone.

The seasonal adjustment policy should reflect the choices of a common revision policy for producing the regular statistics. The revision policy provides the users with the necessary information to cope with revisions. A well-established revision policy defines a predetermined schedule for revisions and is reasonably stable from year to year. The revisions to seasonally adjusted data should be in line with the producer's overall revision policy.

Here we're going to discuss the need to re-design statistical releases taking into account the features of the seasonally adjusted data. The release practices are for each office to design according to its own policy, resources and user demand. In this chapter we merely offer ideas for consideration in developing release practices for seasonally adjusted data.

The users of statistics need adequate support in using seasonally adjusted data. Anyone may read from the newspaper that "the seasonally adjusted industrial production grew in May by 0.5 per cent from April". It tells the reader that production has increased and when it happened, but the reader may not be familiar with the concept of seasonal adjustment. The job of statisticians doesn't end in index calculation since the users will need support in understanding and interpreting statistics.

Documentation

The process of seasonal adjustment should produce several types of documentation: metadata for users, documents needed for production and it should store the resulting data itself. To make use of statistics, the users need to know how statistics have been compiled. The users have different uses for the seasonally adjusted data, and thus they have varying needs related to the levels of metadata. Good documentation not only benefits the users of statistics but also inside the organization it's a prerequisite for sound statistical production.

The ESS Guidelines on Seasonal Adjustment (Eurostat, 2009) pays attention to storing the outputs of seasonal adjustment. At least raw and seasonally adjusted data should be stored within a secure and usable database environment.

Additionally, the database could include the related time series such as calendar adjusted data, trend-cycle and seasonal components. The database would ideally enable comparison of the seasonally adjusted series for each separate time span. For this purpose, it would need to contain separately all the published seasonally adjusted series, the so-called data vintages.

This way the database would be highly useful in revision analysis both for disseminating average revisions and for analysing the appropriateness of the selected seasonal adjustment model and specifications. With archived time series, you can improve the quality of seasonal adjustment in the longer term. It makes it possible to analyse the behaviour of the seasonally adjusted series, for example, during a turning point in the economy.

The purpose of internal documentation is to make it possible to maintain the high quality of seasonal adjustment. Internal documentation comprises the following:

- Step-by-step working instructions for performing seasonal adjustment.
- Internal quality reports describing the quality and special features of the seasonally adjusted and the raw data.
- Regularly updated lists of national holidays if Demetra+ doesn't include them in the pre-defined calendars.
- Lists for monitoring and enumerating the reasons for outliers.

Internal documentation should allow any colleague acquainted with seasonal adjustment to repeat the process. Sufficient documentation helps preserve the skills needed in seasonal adjustment in the organization and is useful for briefing and training others.

Demetra+ presents a considerable amount of quality diagnostics depending on the chosen seasonal adjustment method. By using the XML files generated by Demetra+ you can also share the metadata. For the key series, you can copy the information of the summary statistics of the *Results* panel, i.e. the first page of Main results, Pre-processing, Decomposition and Diagnostics. It will be useful to be able to go back to old quality diagnostics and compare the quality of new results to the previous diagnostics, especially after re-estimating the seasonal model and parameters. More time should be invested in preparing and reading the documentation of the results of key time series.

Every few years, a more comprehensive quality analysis could be undertaken. For regular use, an internal quality report should be simple so that the statistician can easily update it, but it should contain the most relevant details. It should be understandable to colleagues not acquainted with seasonal adjustment, and it could favour charts and tables due to their easy readability. User documentation is usually more concise than the internal quality reports, but it could make use of the same contents. The user documentation is often more fixed, and therefore, quick to update as part of the monthly or quarterly production process.

There are several international guidelines stressing the need to foster transparency of statistical production. *The Fundamental Principles of Official Statistics* (UNECE, 1992) emphasise the importance of transparency in statistical production: "To facilitate a correct interpretation of the data, the statistical agencies are to present information according to scientific standards on the sources, methods and procedures of the statistics."

In the excess of information available in the society, the content of metadata for the users of statistics needs to be well designed. It should include only relevant information. Metadata should enable the users of statistics to understand the idea of seasonal adjustment, use the data, and if needed, replicate the results of seasonal adjustment.

The *Guidelines for the Modelling of Statistical Data and Metadata* (UNECE, 1995) provides a definition for metadata. Metadata provide information on data and about processes of producing and using data. Metadata describe statistical data and - to some extent - processes and tools involved in the production and usage of statistical data.

More specifically, *the ESS Guidelines on Seasonal Adjustment* include recommendations on metadata related to seasonal adjustment. The *Guidelines* include a metadata template that offers ideas for defining the contents of metadata that fits national needs.

First of all, the users need an explanation of seasonal adjustment aimed at the general public, including instructions on how to interpret and use the results. Many statistical offices maintain explanations about their methods and practices on a dedicated part of their website, labelled, for example, "Understanding statistics". The Australian Bureau of Statistics (ABS) offers on their website information on the basics of time series methodologies, including easily understandable explanations to the concept of time series, seasonal effects, components of time series and identifying seasonality. They use the SEASABS software in their seasonal adjustment. It is based on X-11 and X-12-ARIMA. They explain the need for seasonal adjustment in the following way[11]:

"*Seasonal adjustment* is the process of estimating and then removing from a time series influences that are systematic and calendar related. Observed data need to be seasonally adjusted, as seasonal effects can conceal both the true underlying movement in the series as well as certain non-seasonal characteristics which may be of interest to analysts."

In addition to the general explanation of seasonal adjustment, the user documentation should offer details on how seasonal adjustment is performed and include preferably the following information:

- The seasonal adjustment method (e.g. TRAMO/SEATS or X-12-ARIMA) and software (e.g. Demetra+) used.
- General decision rules applied in the process of seasonal adjustment.
- Description of the quality of the raw data.
- Means for outlier detection and correction and information about the events causing outliers in the key time series.
- The choices in calendar adjustment and treatment of national and moving holidays.
- Set of quality indicators for assessing the quality of data.
- The timing and reasons for revisions to the seasonally adjusted data.
- Contact information to the experts.

The general decision rules refer, for example, to application of direct or indirect seasonal adjustment for aggregation of time series. The following section on seasonal adjustment policy will discuss these choices. The main quality indicators used for approving the results of seasonal adjustment are also part of these general decision rules.

The description of the quality of the raw data can simply be included in the usual metadata of statistics. For the purpose of seasonal adjustment, it would be then sufficient to note that the length and the qualities of the series are appropriate for seasonal adjustment. The metadata should also inform the user about any breaks in the series. If generalized based on *the International Recommendations for the Index of Industrial Production*, the metadata of the statistical indicator would already explain the definitions being measured, limitations of use, index compilation methods, weighting system, treatment of changes and departures from international standards (UNSD, 2010).

Understanding outliers helps interpret the seasonally adjusted data. Outliers stay visible in the adjusted data. They contain information about a particular event in the economy. The criteria for identifying outliers and the methods for treating these abrupt changes need an explanation. The producers of seasonally adjusted data could also remind the users of statistics about the events that caused the outliers of the key time series, i.e. strikes. A particular difficulty is the treatment of outliers in the end of the series, where the duration and type of the outlier isn't known.

As recommended by *the Handbook on Data and Metadata Reporting and Presentation*, statistical offices should provide at least a minimum amount of information that would enable assessment of the reliability of each seasonally adjusted series. The statistical offices should maintain sufficient enough metadata to enable users to seasonally adjust, in a consistent way, other series that may not have been seasonally adjusted. (OECD, 2007.)

The United States Bureau of Economic Analysis (BEA) offers a site including frequently asked questions related to seasonal adjustment[12]. The BEA applies X-12-ARIMA in their seasonal adjustment, and they have defined a minimum set of quality diagnostics which they publish together with statistical releases. Among other quality indicators, they publish the overall quality assessment statistics and some indicators which analyse the stability of estimates and the presence of moving seasonality in the key time series. Their explanation about the reasons for revising seasonally adjusted figures includes the following:

11 www.abs.gov.au/websitedbs/d3310114.nsf/4a256353001af3ed4b2562bb00121564/b81ecff00cd36415ca256ce10017de2f!OpenDocument

12 www.census.gov/const/www/faq2.html

"There are two reasons that we revise seasonal factors: We revise factors when we revise the unadjusted data to achieve a better fit to the revised data." and "…when future data become available, we use them to obtain improved seasonal factor estimates for the most recent years of the series. These revised factors lead to revised seasonal adjustments of higher quality."

In addition to the revisions due to changes in the unadjusted original data and the Reg-ARIMA model, revisions are also caused by the two-sided filter and the use of forecasts in seasonal adjustment. In TRAMO/SEATS and X-12-ARIMA the main seasonal adjustment filters are two-sided meaning that the estimator of the seasonal component depends on the observations prior and after a certain period. Thus, the estimator for recent periods requires observations not yet available. First, the preliminary estimators are obtained by ARIMA forecasts and, later, as new observations become available the estimators will be revised until the filter is completed, and the historical estimator is obtained.

To anticipate revisions Demetra+ includes a revision history test of Findley et al (1990) to indicate which series may have excessive revisions. SEATS also provides the standard deviation of the revision in the seasonal and trend-cycle component. When it is excessively high the series may not be worth adjusting.

Usually, statistical offices inform the users in advance about the timetable for revising the seasonal models and parameters. Some do this reanalysis once, some twice a year and some offices more frequently. The average historical revisions of the key economic indicators would provide the user of statistics with a useful tool for anticipating the magnitude of future revisions. For example, the average revisions of the month-on-month changes during the last 24 months give an indication of the expected future revisions. Transparency about the past revisions assists the user in making conclusions based on the data. If no such information is available, the revision can be an unpleasant surprise.

Ensure an easy access to the relevant metadata for the users of statistics. Any release of statistics should include metadata and direct the reader to the more detailed information. It's a widespread practice to publish a link to the metadata in the statistical release. The link may lead to a quality report for the statistical indicator in question and to an archive of the historical quality descriptions.

Seasonal adjustment and revision policy

Seasonal adjustment policy

Many statistical offices need to produce a massive amount of seasonally adjusted data. Often, the substantive units perform seasonal adjustment of the indicators they produce. In some cases, the responsibility of applying seasonal adjustment is vested in one methodology unit. Due to the hectic schedule of statistical production, some degree of decentralisation is necessary in the division of work. This calls for a definition of clear and practical policy for seasonal adjustment.

Defining seasonal adjustment policies to reflect international guidelines would help achieve, gradually, a higher degree of comparability of data between countries. The policy should also reflect the needs of national users of statistics and the resources available in the organization for implementing the policy.

Even with the available modern software, seasonal adjustment is time-consuming. The key statistical aggregates will be in the centre of attention of the national and international users. Consequently, the statistician needs to use more time on the seasonal adjustment of these aggregate time series. For wider use of seasonal adjustment, the organization needs sufficient computer resources, in particular for dissemination of data and storage of the data vintages.

Before starting a large-scale seasonal adjustment, the producer should consider the advantages and disadvantages of seasonal adjustment. Comprehensive testing of seasonal adjustment methods and choices should precede the formulation of the policy. Defining policies to be implemented throughout an entire organization requires consideration and experience.

The policy could be developed in stages, first to cover the basic choices, such as the seasonal adjustment software and method, timing of revisions and guidelines for releasing the adjusted data and its metadata. As experience accumulates, the policy could expand to more detailed instructions for statisticians with problematic time series, breaks in time series and to applying seasonal adjustment at times of economic uncertainty. Similarly, the scope of releasing seasonally adjusted data could be gradually increased.

A seasonal adjustment policy should reflect the knowledge gained in practice and vested in statisticians within the organization, and it should respond to the expectations set by the international guidelines and the users of statistics. The policy should cover at least the following issues:

- A common seasonal adjustment method to be applied.
- Software solutions for seasonal adjustment, dissemination and storage of data.
- Methods and timing of re-analysis, i.e. a revision policy for seasonal adjustment.
- Means of aggregation from lower levels of industrial activity classification to higher levels, or from regional indicators to the country level and in time.
- Treatment of outliers.
- Requirements for the internal documentation and the metadata for users.
- Guidelines for releasing seasonally adjusted data as part of the regular release programme.

Not all statistical offices have decided to select one seasonal adjustment method. Sometimes, for example, the national accounts, industrial, trade and labour statistics could all apply their own practice. However, for clarity and better consistency it could be a good idea to choose one method to be applied in the entire organization bearing in mind the usefulness of regular comparison of the results of alternative methods.

Using commonly applied software and methods upgrades comparability of statistics internationally. To this end, Demetra+ is a useful solution as Eurostat maintains and develops it constantly around the two methods: TRAMO/SEATS and X-12-ARIMA. Releasing seasonally adjusted data may require changes in the dissemination software. The users appreciate access to time series data in such a format that enables the data to be reprocessed without excess manual work. Seasonal adjustment multiplies the number of data cells produced for each statistical indicator which makes dissemination databases a sensible choice for releasing the longer time series. Solid database structures are also needed for storing the results of seasonal adjustment, especially for revision analysis.

Revising the seasonally adjusted data

An essential question in formulating a seasonal adjustment policy is to define the revision practices. As seasonal adjustment is based on estimation, so it's subject to revisions. As explained by the BEA, seasonally adjusted data revise, first of all, due to corrections of raw data and new observations. Second, the refreshed and accumulated data lead to better estimates of the seasonal pattern and to revisions in the Reg-ARIMA model, filters and forecasts. Revisions are welcomed, as they derive from improved raw data and the forecasts used in seasonal adjustment become replaced with new observations based on the raw data. This also influences identifying the seasonal pattern which causes revisions in the historical data as well.

The next charts illustrate the influence of new observations and the refreshment strategy on the seasonally adjusted data. The first chart shows the original series of an industrial production index, and how it has changed in each release of new data during four months (chart 9). The revisions are moderate, but the most recent observation is notably low.

Chart 9
Evolution of an original series

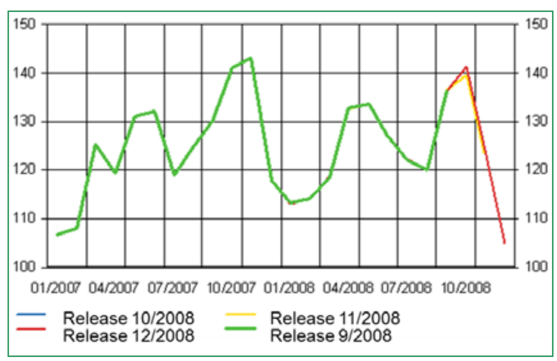
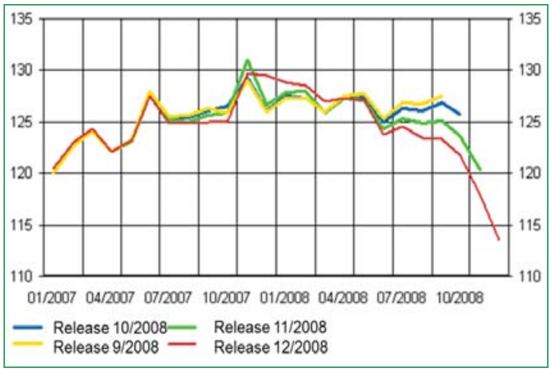


Chart 10 shows the corresponding seasonally adjusted time series. The industrial production index has been seasonally adjusted using partial concurrent adjustment, which is one of the recommended refreshment methods. In other words, the model, filters, outliers and calendar regression variables are re-identified once a year and the respective parameters and factors re-estimated every time new or revised data become available. The seasonal model itself has not been revised during these presented releases.

Chart 10
Evolution of the seasonally adjusted series (without re-identification of the seasonal model)



The chart includes the same four releases as the previous chart. Even though the raw data did not change much at all, the seasonally adjusted time series has changed remarkably with the new observations and the revisions to the parameters of the seasonal adjustment. These revisions took place during the start of a slowdown in the industry which could not be foreseen. Under normal circumstances, revisions would be more moderate in between the re-identification of the seasonal model.

An OECD/Eurostat Task Force has defined general guidelines on revision practices for sub-annual statistics. Countries should establish and maintain a revision policy for producing the regular statistics and have it publicly available. The revision policy would provide the users with the necessary information to cope with revisions. According to the Task Force, a well-established revision policy:

- Defines a predetermined schedule for revisions.
- Is reasonably stable from year to year.
- Is transparent.
- Gives advance notice of larger revisions due to conceptual or methodology changes.
- Offers adequate documentation of revisions.

The recommendation of the Task Force to carry revisions back several years to give consistent time series is important for the quality of seasonal adjustment. Eventually, users will be reassured if they see that revisions take place within the framework of a policy and according to a predetermined schedule. [13]

[13] www.oecd.org/document/21/0,3343,en_2649_34257_40016853_1_1_1_1,00.html

The revision practice as part of the seasonal adjustment policy is worth careful consideration. To identify specific needs of users, consult the main users about the planned revision policy. The policy defines the process for updating the estimates of the published seasonally adjusted data. It consequently defines the frequency of revisions to seasonally adjusted data. However, the intensity of revisions depends on the choices made during seasonal adjustment, e.g. the quality of the forecasts used to extend the series and from the revision to the original data.

The revision policy for seasonal adjustment should address at least the following points:

- Select methods for refreshing the seasonally adjusted data.
- Set the timing for refreshing the adjusted data.
- Define the time period over which the raw and the seasonally adjusted data will be revised.
- Convey the relative size of revisions of the seasonally adjusted data and the main causes of revisions.
- Set the timing of publication of revisions to the seasonally adjusted data and publication of the revisions to the raw data.

In theory, the quality of forecasts used in seasonal adjustment increases with the frequency of updates of the underlying model. There's thus a trade-off between the cost of performing frequent revisions and the quality of data published. However, to complete the equation, the very frequent updates of the seasonal model could also lead to weaker stability of results and revisions in opposing directions.

As we saw in the previous chapters, the current adjustment strategy minimizes the frequency of revision, and the concurrent adjustment strategy generates the most accurate seasonally adjusted data but will lead to more revisions. The seasonal adjustment policy should select between the alternative refreshment strategies. The balanced alternatives between these two extremes may provide better quality of adjustment.

Partial concurrent adjustment is widely used. In its basic form, it keeps the model, filters, outliers and calendar regression variables fixed until the annual or biannual re-identification. It re-identifies the respective parameters and factors every time new or revised data become available.

As we've already seen, Demetra+ offers all together three choices of different partial concurrent adjustment strategies to choose from. The *ESS*

Guidelines (Eurostat, 2009) suggest following one of these balanced approaches between the extremes of current and concurrent adjustment. The choice of the refreshment strategy also depends on the properties of the series.

The optimal frequency of updating the seasonal models is once a year, at least, according to the *ESS Guidelines*. However, if you find a problem between the updates, it should be promptly corrected. There are more degrees of freedom for the frequency of updating the parameters of seasonal models. As a general rule, when the series are shorter than seven years, the specification of the parameters used for pre-treatment and seasonal adjustment could be checked more often, for example twice a year in order to deal with the higher degree of instability of such series. As we've already seen, series shorter than three years shouldn't be seasonally adjusted. Back-recalculated time series are particularly useful to stabilize the seasonal adjustment. User documentation should alert for the greater instability of seasonally adjusted data for relatively short time series.

In the annual re-analysis, you need to examine revisions and balance between accuracy and stability. Stability is also important for the quality of seasonally adjusted data. If TRAMO/SEATS or X-12-ARIMA selects a different model in the annual update, you should examine the diagnostics to find out whether it's notably better than the previous one. Assess also the effect on historical data and check the significance of the regression variables to identify any need for changes in the new specifications.

The refreshment policy needs to consider the time period over which the results are revised. The revisions resulting from the revised seasonal models may be but don't necessarily have to be published in their entirety. Carrying out a full revision from the beginning of the series, promotes a methodically uniform treatment of all values and the easy replication of the seasonally adjusted results. However, some statisticians have questioned whether a newly added figure contains relevant information for significant revisions in the historical seasonal pattern. To balance with the information gain and the revisions, some offices have chosen to limit the period of revision of the seasonally adjusted data to a period that is about four years longer than the revision period for the original data. For the earlier periods, some offices have chosen to keep their seasonal factors unchanged. The information provided by diagnostics – such as the sliding spans and the revision history – may support making the choice of the appropriate revision horizon.

In case of any limits to the revision horizon, this needs to be communicated clearly to the users of statistics. In situations where the raw data are revised from the beginning of the series – for example due to changes in definitions or sampling scheme – the entire seasonally adjusted series should be revised. The consistency between the raw and the seasonally adjusted series should be preserved.

The revision practice, as part of the seasonal adjustment policy, will be based on the analysis of the size of the expected revisions and the effect of the different choices made during seasonal adjustment. It provides information on revisions and their causes to inform the users of statistics. Furthermore, as recommended by the OECD/Eurostat Task Force, the general revision policy introduces a foreseeable framework for revisions.

In general, statistical offices publish revisions to the seasonally adjusted series at the same time, as they add a new month or a quarter of observations to the statistical indicator. They often link a change of the seasonal adjustment method to other methodological revisions of statistics, such as changes of base year or the economic activity classification. Advance information about the forthcoming methodological changes should precede the actual revisions. If mistakes were made additional release to correct the information may be needed. Otherwise, revisions of seasonally adjusted data are usually linked with the regular release schedule of statistics.

The approach to time aggregation should be mentioned in the seasonal adjustment policy. Seasonality is not neutral over the year. It's possible to force the sum, or average, of seasonally adjusted data over each year to equal the sum, or average, of the raw data. It may be difficult for some users of statistics to understand why the annual change differs between the seasonally adjusted and the original data. However, from a theoretical point of view, there's no justification for forcing the annual changes to be equal.

Up to now, only mixed evidence has been provided on the superiority of either direct or indirect aggregation of seasonally adjusted data (ECB, 2003). The direct approach means that the aggregate time series are seasonally adjusted independently, the same way as the so-called lower level component series. The indirect approach refers to deriving the higher levels by aggregating the seasonally adjusted series of the component time series by using a weighting scheme.

The direct approach is preferred for its transparency and accuracy, especially when the component series show similar seasonal patterns. The indirect approach may be preferred when the components of series show seasonal patterns differing in a significant way. Regardless of the simplicity of the direct approach, the indirect approach could be useful in addressing strong user requirements for consistency between lower and higher level aggregates. If you choose indirect seasonal adjustment, the presence of residual seasonality needs to be monitored carefully.

The seasonal adjustment policy should explain how volatility and outliers are treated. It defines the minimum requirements for the user documentation and internal documentation and could suggest templates to be used. The choices for release practices follow the needs of national users and traditions of the office. Improvements could build upon the ideas provided in the next section of this chapter.

Release practices

Time series that seasonal movements have been eliminated from allow comparison of two consecutive months or quarters. Without seasonal adjustment, change should be calculated from the same month of the previous year, i.e. one should compare January to January, February to February etc. However, this comparison isn't free from calendar effects, as the number and assortment of different days in a month vary. It also offers a historical picture of the growth.

Seasonal adjustment makes different months of the series comparable. Thus, it may provide faster indications of changes in the economy, as the changes don't have to be calculated from 12 months back. The choice of growth rates for statistical releases requires careful judgement by the producer of statistics.

Period-to-period growth rates are changes expressed with respect to the previous period, e.g. from April to May. Period-to-period growth rates may often be referred to as month-on-previous-month or month-over-month growth rates, 1-month rate of change, or change on the previous month. Such rates are expressed as $(M_t/M_t\text{-}1)$ -1 or $(Q_t/Q_t\text{-}1)$ -1. M_t denotes the value of a monthly time series in month t and Q_t the value of a quarterly time series in quarter t.

Year-on-year growth rates are rates of change expressed over the same period, month or quarter, of the previous year. They may be referred to as year-over-year growth rates, year-to-year growth rate, rate of change from the previous year, or 12-month rate of change. Such rates are expressed as $(M_t/M_t\text{-}12)$-1 or $(Q_t/Q_t\text{-}4)$-1.

Year-to-date growth rates are data expressed in cumulative terms from the beginning of the year; sometimes referred to as cumulative data. For example, they may compare the sum of values from January 2011 to April 2011, to the same period of 2010.

The above-mentioned growth rates may be calculated for any components of time series, e.g. for the original data, seasonally adjusted data or trend-cycle.[14]

Table 16 has been compiled using different growth rates based on one industrial production index. The underlying data are the same in all columns. The table underlines the difference in timing of detecting turning points by using the different growth rates. The year-to-date growth rates, i.e. the cumulative growth rates, show a turning point in January 2009. The year-on-year growth rate, from the same month one year ago, detects a negative value only one month earlier, in this case. These two provide the slowest indications of changes in the series.

The period-to-period growth rates of the original series don't give any clear indication of turning points. They are influenced by seasonality, as they compare different months of the year which are influenced by seasonality.

The fastest indication of a turning point is given by the period-to-period growth rates of the seasonally adjusted and the trend-cycle series. Yet, the trend-cycle revises slowly and may not show changes this clearly if they occur close to the end of the series. Closer to the end of the series, the seasonally adjusted data would also be somewhat uncertain. Nonetheless, for detecting a turning point, one would in any case need several consecutive observations that give similar signals. Based on the table it is clear that the period-to-period growth rates of the seasonally adjusted series alert the users most swiftly about the turning points: A downward swing occurs already in June 2008, and a slightly better period starts from April 2009.

As enshrined by *the ESS Guidelines on Seasonal Adjustment* (Eurostat, 2009), in all cases, the

[14] For further detail see OECD, 2007.

Table 16
Detecting turning points from different kinds of growth rates

	year-to-date, (cum.) original series	period-to-period, original series	year-on-year, original series	period-to-period, seasonally adjusted	period-to-period, trend-cycle
2008 M01	4.0%	**-7.6%**	4.0%	1.8%	2.4%
2008 M02	7.4%	**-2.0%**	11.1%	5.8%	3.4%
2008 M03	8.3%	8.9%	9.8%	0.4%	1.8%
2008 M04	8.8%	**-1.5%**	10.2%	0.5%	0.6%
2008 M05	10.0%	0.9%	14.8%	0.9%	0.5%
2008 M06	11.1%	**-1.3%**	17.1%	**-0.3%**	**-0.7%**
2008 M07	10.8%	**-2.4%**	9.3%	**-3.1%**	**-1.7%**
2008 M08	10.4%	2.4%	7.4%	**-0.1%**	**-0.8%**
2008 M09	10.1%	**-1.6%**	7.7%	**0.0%**	**-0.4%**
2008 M10	9.8%	4.3%	7.3%	**-1.4%**	**-1.6%**
2008 M11	9.0%	**-3.7%**	1.4%	**-3.4%**	**-2.6%**
2008 M12	8.0%	3.5%	**-1.2%**	**-2.0%**	**-2.1%**
2009 M01	**-4.2%**	**-10.5%**	**-4.2%**	**-0.9%**	**-1.4%**
2009 M02	**-7.6%**	**-9.0%**	**-11.1%**	**-1.5%**	**-1.3%**
2009 M03	**-9.4%**	6.9%	**-12.7%**	**-1.3%**	**-0.9%**
2009 M04	**-10.1%**	**-1.0%**	**-12.3%**	0.7%	0.2%
2009 M05	**-10.6%**	0.7%	**-12.4%**	0.6%	0.6%
2009 M06	**-9.4%**	8.5%	**-3.7%**	9.1%	0.3%
2009 M07	**-8.7%**	**-2.6%**	**-4.0%**	**-3.0%**	**-0.2%**
2009 M08	**-8.3%**	0.7%	**-5.5%**	**-1.9%**	0.9%
2009 M09	**-7.6%**	1.6%	**-2.4%**	3.5%	2.4%
2009 M10	**-6.9%**	6.4%	**-0.5%**	0.0%	1.6%
2009 M11	**-6.0%**	**-0.8%**	2.5%	**-0.3%**	0.8%
2009 M12	**-5.0%**	7.6%	6.5%	1.6%	0.7%
	slowest	not informative	slow	fast	fast, end problem

information contained within the statistical release should adhere to the principles of ensuring transparency and assisting users in making informed decisions. One should ensure sufficient resources and enough time for analysing the results of seasonal adjustment before publishing the data for the first time. Introducing seasonal adjustment requires re-structuring the content of the website to facilitate easy access to time series data. It requires time to re-design the statistical news releases so that they would support the users.

The challenge with statistical releases is to offer the relevant information in a concise form while keeping the message simple and understandable. At the same time, the release should contain somewhat comprehensive information so that the users can draw correct conclusions. To help statistical offices with this challenge, *the Handbook on Data and Metadata Reporting and Presentation* comprises a set of practical recommendations on releasing seasonally adjusted data. The recommendations have been drafted in international cooperation based on the experience of a number of short-term statisticians. The following text draws on those recommendations, and Annex 1 provides a summary of the recommendations extracted from the *Handbook* published by OECD (Annex 1).

When seasonality is present and identifiable, seasonal adjustment would improve the readability of statistics. In these cases, the statistical office should release the data in the seasonally adjusted form. The producer can choose the level of detail for providing seasonally adjusted data by taking into account the user demand and the available resources. They can increase the amount of available seasonally adjusted data in the course of time.

If seasonality influences the indicator ideally the statistical release would focus on the seasonally adjusted data. Users should also have access to the original series, either in the publication or by reference to it. The original series contains all characteristics of the data. The seasonally adjusted data contain the news of the series by combining the trend-cycle and the irregular component. As press releases aim to provide news, seasonally adjusted data are the appropriate kind of data to be presented. Where there's a user demand, the producer may also disseminate components of seasonal adjustment, e.g. the calendar day adjusted and/or the trend-cycle series.

Statistical releases for data with seasonal influences should at least provide period-to-period growth rates for the latest period. If the statistical tradition includes publishing levels of the measured variable, and if space permits, the change or the value in levels may also be published in the release. Cumulative growth rates may be useful as additional information, but not as the main focus.

The year-on-year growth rate is not always the same for the original and the seasonally adjusted data, unless the seasonal pattern is stable. As we discussed previously, seasonality may evolve over time as reflected in the seasonally adjusted series.

A number of statistical offices release both the change from the previous period based on the seasonally adjusted data, and the change from the same period of the previous year based on the calendar adjusted or original data. This is also recommended by *the International Recommendations for the Index of Industrial Production* (UNSD, 2010). The following extract illustrates how Statistics Sweden (SCB) releases their industrial production index (image 73)[15].

Notably, the release of SCB says that the latest figures are preliminary and mentions the main revisions to the previous release. The related metadata should include the explanations of the main sources of revisions and explain any other quality issues such as breaks in the series.

The seasonally adjusted data are the best way of presenting period-to-previous-period changes, even if the irregular component which belongs to the seasonally adjusted data, is relatively large (OECD, 2007). For highly volatile seasonally adjusted series, the period-to-period changes may change direction frequently. It's, therefore, useful to analyse the behaviour of these changes to see the kind of message they would deliver to the reader. The statistical releases focus on the key economic indicators which are usually higher level aggregates and less volatile than individual economic activities.

Image 73
An extract of a statistical release

Industrial production index (IPI), July 2011:

Increase in industrial production in July

Industrial production increased by 2.8 percent in seasonally adjusted figures in July compared to June. Comparing July with the same month of the previous year, industrial production increased by 8.2 percent in working day adjusted figures.

The production in the industry sub-sectors developed in different directions in July compared to June. The motor vehicle industry had the strongest development with an increase of 22.3 percent.

The development was positive when comparing the latest three month period, May-July, with the previous three month period. Production within industry increased by 1.1 percent.

Industrial production continued to develop positively on a yearly basis, but like last month, less strongly than earlier this year. Production within industry increased with 8.2 percent in July compared with the same month of the previous year. As previous months, development was the strongest in the electrical equipment industry with an increase of 39.9 percent.

The figures for July are preliminary. Since the previous publication the change in industrial production in June compared to May has been revised upwards by 0.5 percentage points amounting to a decrease of 2.8 percent. The change in the production in June compared to the same month of the previous year has been revised upwards by 0.6 percentage points amounting to an increase of 7.0 percent.

Image 74
An extract from statistical news

The seasonally adjusted number of visits to and from the UK rose in the three months May to July 2011 when compared with the previous three months. However, visits abroad completed in July alone were down compared with July 2010.

[15] www.scb.se/Pages/PressRelease____320267.aspx

To deal with high irregular movements that may blur the trend of the series, the statistical release may also concentrate on a rate of change based on the sum of three months or a quarter (image 74).

Image 74 is an extract from the website of the Office for National Statistics (ONS) of the United Kingdom. In the news about monthly tourism statistics, they focus on the three months' sum of the seasonally adjusted series and compare it with the previous three months. In the statistical release, they offer more detailed information and the main contributors to the development.[16]

The rate of change compared with the same period of the previous year, i.e. the year-on-year changes should be calculated based on the calendar day adjusted series, or if unavailable, based on the original series. The calendar day adjusted series, i.e. the working day or trading day adjusted data, make the same months of different years more comparable by correcting the variation caused by the number of working days etc. Any special events in the previous year affect the year-on-year changes. Where necessary, the reader should be reminded of these effects when presenting year-on-year changes. The example related to moving holidays in chapter 2 illustrates clearly the utmost importance of noting these special events. The statistical release of Statistics Finland from 2008 provided the reader with alternative growth rates and mentioned the difference in working days compared with April 2007 (image 75)[17].

Statistical releases can only contain a limited amount of information. Nonetheless, the main contributors to change are interesting to the users of statistics. Sometimes, the aggregate growth rate may remain the same regardless of many changes in the development of the sub-populations. Statistical offices often have the sole access to large data sets and could inform the user of the variety of important changes in the sub-populations.

The International Recommendations for the Index of Industrial Production (UNSD, 2010) recommend presenting the development of those product groups or industries that are primarily responsible for the monthly movement in the aggregate index.

Seasonally adjusted data are at their most useful form in a time series format. Users of statistics need an access, preferably electronically, to the complete time series. The website should offer the original and the seasonally adjusted data in their full length. In addition, the users may prefer to have access to the trend-cycle and the calendar day adjusted time series. Index numbers, in a time series form, enable comparison of different sectors of the economy or different regions or countries more easily than levels or individual change percentages. It is not sufficient to provide the users of statistics only with the latest growth rate without any time series.

The users need to be notified of the so-called end-point problem of the trend-cycle series. Significant revisions may occur in the end of the trend-cycle series mainly due to the accumulating new observations. Turning points can be identified reliably often only after several months. Therefore, the trend-cycle series of the most recent observations should be presented to users with caution, as they are uncertain and can suffer of phase-shift problems.

A real data example of the trend-cycle series illustrates the end-point problem (chart 11). In May 2008, the end of the trend-cycle series depicts an upward development in the start of 2008, whereas ten months later, in March 2009, the trend-cycle displays a turning point, a downward swing from the start of 2008.

Image 75
An extract from a statistical release related to the Easter effect

Published: 8 July 2008

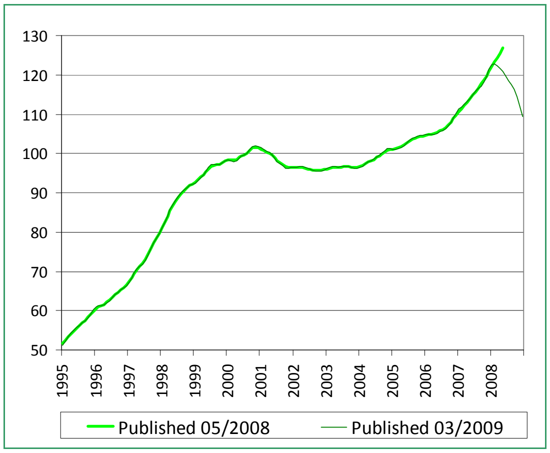
Output of the national economy grew in April

Seasonally adjusted output rose by 1.2 per cent in April from the month before. Year-on-year the increase amounted to 8.1 per cent according to the original series. April 2008 had tree working days more than the comparison month of the previous year. Adjusted for working days, the year-on-year growth was 5.1 per cent. The data derive from Statistics Finland's Trend Indicator of Output.

16 www.ons.gov.uk/ons/dcp171778_233168.pdf

17 www.stat.fi/til/ktkk/2008/04/ktkk_2008_04_2008-07-08_tie_001_en.html

Chart 11
End-point problem in the trend-cycle



The ESS Guidelines on Seasonal Adjustment (Eurostat, 2009) don't recommend showing the most recent values of the trend-cycle estimates because of the end-point problem. In addition, the producers should publish information on average revisions of at least the seasonally adjusted series.

Several international guidelines give tips for re-designing the process and content of statistical releases. For example, *the International Recommendations for the Index of Industrial Production* (UNSD, 2010) offer a set of general recommendations for releasing industrial production indices. Among other things, it adds that data should be released as soon as possible but notes the trade-off between timeliness and quality. The releases should follow an advance release calendar and consistent release practices. The statistical offices should make data available to all users at the same time and foster the confidentiality of individual survey respondents. Providing the users with the contact details of relevant statisticians who can answer various questions would be a good practice.

User support

The statistical offices are facing a growing demand for timely statistical information on the economic development. At the same time, an increasing information overflow surrounds the users of statistics, and rapid development of new communication tools changes the way information is exchanged. Releasing statistical information doesn't occur in isolation. There are distinct differences between the users of statistics, in their technical abilities and understanding of statistics. This brings about challenges for supporting the users in making correct conclusions.

An everyday user of statistics benefits from the availability of seasonally adjusted data. As discussed, it makes it possible to compare development in different industries and countries and could offer faster indications of changes in the economy. Some users may re-process the results further in their own work. For example, economists and analysts base their forecasting models on the seasonally adjusted data. Sometimes, they may perform their own seasonal adjustment with their own specifications. Therefore, the data should be in an easily accessible format, and additional data, such as different components of the original data, could be made available at request.

As mentioned in the previous text, user support is based on the transparency of metadata, availability of general explanations on the statistical methods applied and on the good design of statistical releases and websites. An effective news release is one that (UNECE, 2009):

- Tells a story about the data.
- Has relevance for the public.
- Catches the reader's attention quickly with a headline or image.
- Is easily understood, interesting and often entertaining.
- Encourages others, including the media, to use statistics appropriately to add impact to what they are communicating.

The Internet provides increasing possibilities for disseminating statistical data, but at the same time it consists of an excess of data with different qualities. The Internet is an efficient data dissemination tool, but it requires efforts from the data producers to keep track of the unknown users of statistics in order to understand their needs and problems in using statistical data. The first section of this chapter discussed documentation issues related to seasonal adjustment. It mentioned some examples of statistical offices having created a website for helping the users in "understanding statistics". These sites often comprise a set of frequently asked questions that provide useful user support.

The society has developed a snack culture in information consumption (UNECE, 2009). This has already influenced the statistical releases used in delivering the news in statistics. The releases are concise and start from the most important message they wish to convey. Therefore, the releases usually start with the most recent information, based on the seasonally adjusted data. Nowadays, the user needs to know quickly the main conclusions for immediate processing.

Prior to releasing seasonally adjusted data, the staff needs to be ready to explain the procedure and guide the users in the correct use and interpretation of data. One should allocate sufficient resources and time into examining seasonal adjustment and train staff for the purpose of providing the users with the guidance needed for better interpretation of statistics.

The release would ideally provide the contact information for the users of statistics to obtain further information about concepts, definitions and statistical methodologies. The contact could be a generic contact point that directs the requests to the experts. An expert responsible for the released data could be available to reply to users' questions on the latest statistical releases in order to share their expertise. The responsibility of statistical offices doesn't end in the release, on the contrary, according to *the Fundamental Principles of Official Statistics* (UNECE, 1992) statistical agencies are entitled to comment on erroneous interpretation and misuse of statistics.

The producer of statistics undoubtedly benefits from consulting the users of statistics. With the automated dissemination procedures, user-producer interaction becomes rarer. To fill in this gap, many statistical offices carry out stakeholder analyses, conduct user surveys and organise seminars and conferences to discuss statistics with the main users. This kind of interaction helps the producer stay relevant by learning from user needs, and the user to understand the work and methods applied by the producer. Customer training services on analysing economic development by using statistics could include tips for understanding and using the seasonally adjusted data.

Regardless of the complexity of seasonal adjustment it aims at better user service. It should help discover the underlying trend and turning points which are of interest to the users of statistics. The job of the statistician is to provide the best possible data, inform the users about the limits of the data and support them in using and interpreting the information.

Annex 1
Recommendations of the data and metadata reporting and presentation handbook

For quick reference, the data and metadata reporting and presentation recommendations provided by *the Data and Metadata Reporting and Presentation Handbook* (OECD, 2007) are presented below. The numbering refers to the body of the OECD Handbook and has been included to facilitate user reference to the complete text in order to obtain background information behind each guideline, etc. For further information, extensive citations are provided in the Handbook (refer References - Section 8), where in most instances web links have been provided to the entire source document.

The Glossary in the Annex of the OECD Handbook contains definitions of key terms and concepts used in the recommendations. Source information is also provided for each entry in the Glossary.

2.1.2 Recommended practices to ensure consistency in presentation and reporting practices included in future international statistical guidelines and recommendations

It is recommended that the authors of future international statistical guidelines and recommendations for various statistical domains take a more modular approach in the preparation of those standards by using, as required, existing international recommendations in key areas such as the reporting and presentation practices outlined in this Handbook and terminology and definitions presented international glossaries such as the Metadata Common Vocabulary (MCV) described in Section 6.3.4. Such use may take the form of either the direct incorporation of text within future guidelines or by reference / link to the recommendations in this Handbook, etc.

2.2.1 National and international publishing manuals and guidelines

Recommended practices for publication manuals:

- In the interests ensuring consistency in the presentation of data disseminated across different paper publications and electronic media within an organisation, it is recommended that each statistical agency compile a publishing manual or set of guidelines providing guidance to author areas within the organisation.
- It is also recommended that statistical agencies place such publication manuals and guidelines in the public domain by locating them on their websites. This would give greater transparency and external scrutiny to internal practices and provide a means of disseminating recommended practices to organisations in other parts of the national statistical system.

3.5 Recommended practices to ensure the use of consistent terminology and definitions

The recommended practices outlined below are primarily designed to minimise the current common practice of different author areas within national and international organisations developing their own, often inconsistent, sets of terms and related definitions. Even where differences in terminology are appropriate there's still a need to provide users with information about the context / reasons for such differences.

Agencies should therefore:

- Establish a terminology management strategy and associated structures appropriate to their needs, requirements and resource capacity to reduce the use of inconsistent terminology and related concepts by the different author areas in the organisation which in turn are applied in various questionnaires and disseminated output.

There are a range of options available to achieve this objective. These include the creation of a corporate thesaurus which can be readily accessed by different parts of the organisation and in which existing inconsistent definitions for the same term are confronted and differences either eliminated or explained. Where possible, definitions applied at the national level should align with international definitions, modified as appropriate to provide further elaboration and / or to meet national circumstances.

- Irrespective of the tool(s) adopted (glossaries, thesauri), there's still the need for senior management within an organisation to ensure that appropriate practices and principles involving the use consistent terminology are developed and adopted across the organisation.

- Ensure that any thesaurus developed by an agency contains the minimum structures outlined in existing international guidelines such as those formulated by the ISO, OMG and W3C.

Corporate glossaries should contain the following in order to facilitate their interoperability: concept label; definition; detailed source information about where the definition was derived; related terms; and context field providing additional information or highlighting how a definition is used within one statistical domain or geographical context. Finally, the glossary should also be translated into various languages used by the institution in question.

- Provide appropriate cross references or links in domain specific glossaries, on Internet or intranet sites to general glossary databases that have been developed either at the national or international levels, the purpose being to make existing standard terms and definitions more readily available.
- Avoid attaching precisely the same label or title for different definitions (for further discussion refer Section 6.3.5).

4. Guidelines for the reporting of different types of statistical data

4.2.1 For original data

- The availability of original data affords maximum flexibility for users interested in undertaking further analysis or transformation beyond how it is presented in the data source on hand. Therefore, users should have access to at least some of the key aggregates of original data where they exist for a particular series, either directly in the publication or through the provision of references or hyperlinks. The provisions under which access to detailed original, confidentialised data is given would be governed by the organisation's dissemination policy.
- In situations where original data are known to have significant non-sampling errors, appropriate metadata should be provided to facilitate appropriate use of the data. In such instances consideration should also be given to the presentation and use of other types of data such as percentages in any analyses of main features.
- Original data are frequently transformed into growth rates or indices, for example, in a press release to facilitate interpretation and understanding. Ideally original data should be presented in addition to the transformed series. However, this may not be possible in some dissemination media due to space considerations and in this situation it would be sufficient to provide a clear indication of their availability and how they could be accessed.

4.3.2 For indices

The Statistics Canada Policy on Informing Users of Data Quality and Methodology (Statistics Canada 2000, Section E.3) states that the provision of an adequate description of characteristics and methodologies specific to indices is as important to users as quality assessments of the data. Statistics Canada therefore recommends the provision of the following metadata:

- precise definitions of the underlying economic concepts the indices are intended to measure. Specific mention should be given to any limitations in the use or application of the index, citing the example of deflation of macro-economic aggregates; and
- descriptions of the methodologies used in the compilation of the index, with particular reference to the:
    - index calculation methods entailing the choice of index formula (e.g. Laspeyres, Paasche, Fisher) and the strategy for constructing the index series (i.e. as either fixed base or chain indices);
    - weighting system used, weight revision practices and frequency of weight revision;
    - computation at various aggregation levels; o selection of base year;
    - frequency of re-basing; o procedures for linking indices;
    - treatment of changes in the composition of commodities in the market as well as changes in quality.

The methodologies applied should be compared with underlying index concepts and the impact of departures described.

Finally, as much of the above information is of specific interest to specialised users, consideration should be given to having differing levels of information targeted to different kinds of users. The guidelines for the reporting and dissemination of metadata provided in Section 6 emphasise the desirability of structuring metadata appropriately for users with differing degrees of expertise and need.

4.4.3 For growth rates

The two key recommendations with respect to growth rate terminology entail the need for statistical agencies to:

- minimise the risk of user misunderstanding of the growth rates being applied through the consistent use of growth rate terminology and the provision of definitions such as those provided in Section 4.4.2 in accompanying metadata; and
- develop a consistent (and unambiguous) set of expressions for use in analytical text to describe changes in annual and infra-annual growth rates and to apply those expressions consistently both within the same press releases and in other press releases disseminated by the same organisation.

Recommendations for the presentation of growth rates in the context of different forms of data used in time series analyses are summarized in Section 4.6.

4.5.2 For ratios, proportions, percentages and rates

The main issues for the presentation of and ratios centre around the provision of appropriate methodological information (metadata) describing both the actual rates / ratios and, the component series used in their derivation. Precision is required in the provision of information for the user about the time period referred to, the nature of the population being described and the type of occurrence being measured (Palmore and Gardner 1994). More specifically:

- the term "rate", "percentage", or "ratio" should be included in the actual label, e.g. maternal mortality rate, crude death rate, etc., to ensure user understanding that the original data has been transformed;
- for rates where the total population is expressed as unity the unit of measurement used in the population should be included in the table heading for the rate, e.g. per 1 000 live births;
- the series labels in the components that make up the rate or ratio should be based where possible on existing international terminology. These are generally outlined in the international guidelines and recommendations for the relevant statistical domain (refer UNSD 2002a) or glossary databases disseminated by international organisations such the OECD's Glossary of Statistical Terms (OECD 2002a); the Eurostat Concepts and Definitions Database (CODED) (Eurostat 2006a), or UNSD's Definitions for United Nations Common Database (UNSD 2002b). Departures from international concepts should be documented in the metadata accompanying the rate / ratio;
- detailed information about the source(s) of the component series used in the derivation of the rate or ratio should be provided. Minimum information comprises: type of data source (administrative, household survey or census, business survey or census), reference period, full official title of the series, full name of the source agency or institution;
- users should have access to the original data used in the derivation of the rate / ratio. This could be included in the body of the publication where the rates or ratios are disseminated (e.g. as annex tables) or through the provision of sufficient reference information or hyperlinks that will enable users to access the original data;
- users should be provided with information on methodologies used in the compilation of the component series used in the derivation of the rate / ratio. The quality (in particular, comparability both over time and between countries) of the rate or ratio is only as good as the quality of the series used in its calculation, and appropriate metadata is therefore essential to enable users to form an understanding of quality and relevance of the rate /ratio for a particular need or purpose.

The final issue involves the need to standardise extraneous variables used in the compilation of ratios, proportions, rates, etc. As mentioned in Section 4.5.1, such standardisation is necessary to enable the comparison of ratios, etc. between countries, regions, etc., and in some instances over time in the same geographic area.

5.6 Recommended practices for the reporting of different forms of time series data

Forms of time series data to be presented

Recommendation 1: When seasonality is present and can be identified, sub-annual indicators should be made available in seasonally adjusted form. The level of detail of indicators to be adjusted should be chosen taking into account user demand and cost-effectiveness criteria. The adjustment should be applied appropriately using the method chosen as a standard by the agency. The method used should be explicitly mentioned in metadata accompanying the series.

Recommendation 2: When applicable, the focus of press releases (or similar releases to the general public) concerning the main sub-annual indicators should be on their appropriately seasonally adjusted version. Users should also be given access to the original (or raw) series, either in the publication (if space permits) or by reference to it. Where there is a user demand, the agency may also disseminate intermediate components of the seasonal adjustment process (e.g. series adjusted for calendar effects) and / or trend-cycle estimates but it should be clearly indicated that the focus is on the seasonally adjusted estimate when short-term variation is of interest.

Analytical transformations

Recommendation 3: Press releases presenting seasonally adjusted flow series should at the minimum provide period-to-period growth rates for the latest period and, if space permits, period-to-period (e.g. month-on-previous-month, quarter-on-previous-quarter) change in levels.

Recommendation 4: For month-on-previous-month and quarter-on-previous-quarter rates of change, seasonally adjusted data is the best way of presenting information about a time series (trend-cycle and irregular movements) and for presenting short-term developments, even if the irregular component is relatively large. To deal with irregular movements that blur the trend the rate of change bases on two or three months' (or quarters) worth of values can be utilised.

Recommendation 5: For rate of change with respect to the same period of previous year the year-onyear changes should be applied to raw data and to data adjusted for calendar effects if the latter are available. Where necessary, special effects contained in the base period should be highlighted when presenting YoY (base effect).

Recommendation 6: Because of the risk of providing misleading signals, especially where series display significant volatility, the presentation of annualised level changes is not recommended, especially as the key headline series. Where annualised changes are used, users should be provided with information regarding the possibility of misleading signals due to series volatility.

Recommendation 7: Annualised period-to-period growth rates are not recommended for the presentation of quarterly or monthly growth rates. Preference should be given to the use of year-on-year growth rates.

Information about seasonal adjustment to be provided to users

Recommendation 8: Statistical agencies should disseminate a non-technical explanation of seasonal adjustment and its interpretation for the benefit of, and aimed at, the general public.

Recommendation 9: For the benefit of users requiring information about the appropriateness of the seasonal adjustment method applied, statistical agencies should provide a minimum amount of information that would enable an assessment of the reliability of each seasonally adjusted series.

Recommendation 10: Statistical agencies should maintain metadata on seasonal adjustment of sufficient extent to enable outside users to seasonally adjust in a consistent way other series from the same statistical program that may not have been seasonally adjusted.

6.3 Recommended practices for the reporting and dissemination of metadata

6.3.1 The need for metadata

All statistical agencies should:

- compile metadata required for users to understand the strengths and limitations of the statistics it describes; and
- keep their metadata up-to-date, incorporating the latest changes in definitions, classifications and methodology, etc.

6.3.2 Access to metadata

Key recommendations in this area include:

- ensuring that users have ready access to such metadata through its dissemination via a range of different media – paper publications, CD-ROMs, etc. However, it is important for all metadata to be available to users on the Internet, given that it provides the most accessible medium for obtaining the most up-to-date metadata. It is also recommended practice for metadata to be structured in such a way as to meet the needs of a range of users with different requirements and/or statistical expertise. This doesn't necessarily entail the physical presentation of different metadata to each group of users with different statistical expertise. However, a layered presentation of metadata is recommended, progressing from summary metadata to more detailed metadata. Each layer should use clear and precise text;

- dissemination of metadata free of charge on the Internet. There is strong support for the notion that metadata describing statistics has a high public good component and should therefore be disseminated free of charge on the Internet even if the actual economic and social statistics they describe and paper publication versions of the metadata are subject to an organisation's price regime;
- active linkage of metadata to the statistical tables and graphs they describe and vice versa;
- the availability of metadata not only in the national language but also, where resources permit, in a common language such as English;
- structuring the metadata for different statistical domains on the basis of some hierarchic classification. Consideration could be given to the adoption of the UNECE Classification of International Statistical Activities as the international standard for metadata. The September 2005 version is available at www.unece.org/stats/documents/ece/ces/bur/2005/5.e.pdf;
- provision of a local search engine based on free text search;
- implementing recommended practice for ensuring either the stability of URLs (Uniform Resource Locators) or providing links between the old and new URLs that will redirect users to the new address. This is a key issue given the importance of links between websites;
- providing the names of contact persons or email addresses where further information about concepts, definitions and statistical methodologies may be obtained. In some organisations the "contact" would be a generic corporate contact point or referral service for all client enquiries.

6.3.3. Adoption of a set of common metadata items

International agencies should work with national statistical agencies to develop a core set of nondomain specific metadata items (or prompt points) such as those being developed under the Statistical Data and Metadata Exchange (SDMX) initiative – refer para. 175.

6.3.4 Adoption of a common set of terminology for metadata preparation

Considerable resources are often expended by international organisations in verifying text, etc., to ensure that methodological descriptions are as consistent as possible between countries. A mechanism for achieving this would be the rigorous use of terminology imbedded in the various international statistical guidelines and recommendations. This could be facilitated by the use of glossaries published by international organisations which contain definitions derived from those standards, in particular, the Metadata Common Vocabulary (MCV) developed under the umbrella of the SDMX initiative

6.3.5 Unambiguous presentation of similar but not identical statistical data

Five broad recommendations of good practice in this area comprise (Friez 2003):

- Similar but different series should be given different titles to facilitate clear differentiation by users.
- International organisations that disseminate national data should always be aware of and clearly state in their metadata whether or not the precise series they disseminate that are derived from national sources are also disseminated in the country of origin, or compiled and / or transformed by national agencies specifically to meet the requirements of international organisations.
- International organisations should clearly describe in their metadata, specific details of any transformation of national data they perform to make the series more internationally comparable. Data transformed by international organisations should be clearly indicated as such, particularly, but not only where, published alongside different national series for the same statistical domain. The two sets of series must be clearly differentiated in the mind of the user.
- The precise name of the classification used in statistics disseminated by national agencies and international organisations (especially when transformed to an international classification to enhance international comparability) should always be clearly indicated (for instance, NACE Rev. 1, CITI, Main Industrial Grouping (MIG) or national classification) so that when the same denomination is used in various classifications such as intermediate goods, the user clearly knows which classification has been used.
- When a field of activity is only partially covered (such as MIG-intermediate goods or MIGconsumer goods in the new orders

indicators of the European Commission's Short-term Statistics Regulation), it should be clearly indicated for instance with an asterisk or a footnote (for example, in the Eurostat's short-term statistics new orders series, MIG-non durable goods (1) - (1) Partial ; doesn't include NACE 151-155, 158, 159, 16, 19, 22, 364-366).

7.1.3 Recommended practices for data revision

7.1.3.1 Consultations with users elicit views about revisions practices

Preliminary to elaborating a country's revisions policy, it is important to consult the main users of official statistics to identify needs and priorities specific to individual countries. Their views could be sought, for example, about their particular needs for timeliness of data, problems they experience because of revisions, and their priorities about balancing timeliness with accuracy and consistency.

7.1.3.2 Provision of a clear, short readily accessible summary statements to users of when to expect revisions and why revisions are undertaken

7.1.3.3 Where possible establish a stable revisions cycle from year to year.

7.1.3.4 Balance stability of a time series against the need to revise series because of the need introduce new methodologies, concepts, classifications, etc.

7.1.3.5 Ensure that revisions are carried back several years to give consistent time series

To maintain the serviceability of data following major revisions, data should be revised back as far as is reasonable based on a balancing of user needs, costs, and availability of source data. The revised time series should be released simultaneously with the revised current data or soon thereafter, preferably in easily accessible electronic format. The revised series should be of sufficient detail and not so aggregated that users are not able to detect the sources of the changes.

7.1.3.6 Provide appropriate documentation on revisions

Such documentation should include:

- clear identification of preliminary (or provisional) data and revised data;
- provision of advance notice of major changes in concepts, definitions, and classification and in statistical methods;
- information on the sources of revision are explained when the revised series are released;
- information on breaks in series when consistent series cannot be constructed;
- information on the size of possible future revisions based on past history;

7.1.3.7 Users are reminded of the size of the likely revisions based on past history

Users should be provided with information sufficient for them to make an informed judgment about the reliability and accuracy of preliminary or provisional statistics. The following two recommended practices for revision studies have been identified:

- periodic analyses of revisions investigate the sources of revision from earlier estimates and statistical measures of the revisions;
- the analyses are published for major aggregates to facilitate assessment of the reliability of the preliminary estimates.

7.1.3.8 Transparent and timely reporting of errors and mistakes

7.2.3 Recommended practices on the presentation of series breaks

Recommended practice with regards to time series breaks entails:

- The compiling agency taking all possible steps to avoid and minimise changes to questionnaires, definitions and classifications used to collect and compile data. Methodologies should be developed to reduce the frequency of revisions.

However, there comes a time when the time series may be disrupted even when outdated classifications, concepts and questionnaires are maintained. In such instances a complete break in series may be preferred to series that continue to be collected on the basis of outmoded classifications and concepts that don't approximate reality. There is clearly a tradeoff between costs imposed by breaking a time series on one hand and the benefits from improving the relevance of the time series on the other (BEA 1993).

- Where significant breaks in a time series are unavoidable, users should be given warning well in advance of the implementation of the series break outlining the timing of implementation and a detailed explanation of the reason(s) for the change. "In advance" is taken to mean not just the time of implemention but sufficient time to enable users to implement modifications to their systems, programmes or databases and to seek further clarification if necessary. A common practice adopted by many statistical agencies is to issue a detailed discussion paper many months in advance of the change.

- Actual breaks in the series should be clearly identified in both the statistical table and any accompanying graphs. A variety of methods are commonly used by national agencies and international organisations to highlight in tables that a series break has actually ocurred. These include the insertion of a line in the table at the break point, inclusion of a footnote or tabular presentation as an entirely new series. Whichever method is adopted, the main point is that the break is completely clear to users. Consideration will also need to be given to the identification of series breaks (together with appropriate explanatory information) in data disseminated electronically such as via on-line databases, etc.

The following information drawn directly from Eurostat guidelines should also be provided (Eurostat 2003c, p. 16):

  - the reference period of the survey where the break occurred;
  - whether or not the difference reported is one-off with limited implications for the time series and / or if the reported change led to harmonisation with any standards;
  - a precise outline of the difference in concepts and methods of measurement before and after the series break;
  - a description of the cause(s) of the difference, e.g. changes in classification, in statistical methodology, statistical population, methods of data transformation, concepts, administrative procedures with regard to statistical data from administrative sources;
  - an assessment of the magnitude of the effect of the change, where possible, with a quantitative measure.

Links and references to more detailed information should also be provided.

- Points in line graphs shouldn't be joined across discontinuities in data. The reason for the series break should be explained in a footnote accompanying the graph with appropriate links or references to more detailed explanations of the causes of the breaks.
- When methodological changes are introduced, an attempt should be made to revise the historical series as far back as data and available resources permit. Ideally, such backcasting should extend back 2-3 years to reflect the new methodology, etc.

7.3 Presentation of information on sampling and non-sampling errors

7.3.1 Sampling errors

In the interests of data transparency, and to help ensure the appropriate use of data, statistics derived from all sample surveys should be accompanied by information on sampling errors. Such information should be provided for all dissemination media – online databases, websites, other electronic products, paper publications and press releases. It is also important for the information to be expressed in nontechnical terms capable of being understood by the non-specialist user. The mode of presentation and the amount of detail provided should therefore meet the specific needs of particular categories of users (UNSD 1993, p. 176).

The required information comprises the provision of the following information in accompanying or clearly linked technical notes outlining (OMB 2001, p. 3-8):

- Alerting users to the fact that data are derived from a random or non-random sample. If the latter then inference implications should be clearly stated.
- Sampling error should be identified as a source of error which should be explained and interpreted for data users through provision of a brief definition of sampling error. For example, strong warnings about the unreliability of data with high sampling error.
- Sampling errors must be presented in the context of total survey error. In this context users should be made aware of the fact that sampling error is just one, and often not the most significant, component of total error (UNSD 1993, p. 176, 7.1 (1)).
- If statistical tests are used in the report, the significance level at which statistical tests are conducted should be stated explicitly.
- Sampling errors for key estimates should be available to the user either in a table in the publication or linked on the Internet. Some

form of notation should also be placed directly beside estimates with very high sampling (or non-sampling) error. Sampling errors may be presented in one of a number of different forms, for example:

- as absolute values of the standard error (se);
- as relative values, standard error divided by the estimate (rse); or o in the form of probability or confidence intervals.

The preferred use of either the absolute or relative forms depends on the nature of the estimate and readers are referred to the United Nations publication, Sampling Errors in Household Surveys (UNSD 1993, p.178) for a detailed evaluation of the different forms of presentation and several examples of recommended practice. The UN evaluation emphasises the importance of ensuring that the chosen method is clearly and unambiguously described and presented with accompanying definitions and notation.

In order to ensure consistency in the dissemination of this information across the organisation in all published output subject to sampling error, some statistical agencies mandate a standard set of words to be included in all relevant publications.

Where space considerations preclude the inclusion of detailed information, either references or hyperlinks to more detailed technical reports or user manuals should be provided. Such information should enable specialist users to analyse detailed data or compile new tabulations and would therefore:

- identify the specific method used for calculating the sampling error;
- provide sampling error calculations (tabulations) for different types of estimates (e.g. levels, percents, ratios, movements, means and medians) for a number of variables and disaggregations. The aim is to provide a basis for extrapolation to statistics for which sampling errors have not been computed by the source agency (UNSD 1993, p. 180);
- contain evaluations of the procedures used for estimating sampling errors.

7.3.2 Non-sampling errors

The focus of the recommended practice on non-sampling errors outlined below is not the methods by which national agencies minimise their impact but rather guidelines on the type of information on such errors to be reported with disseminated statistics:

As for the reporting of information on sampling error for all sample surveys, all statistical output disseminated by national agencies and international organisations should be also accompanied by information on non-sampling errors. Such information should be accessible for statistics disseminated on all types of media – online databases, websites, other electronic products, paper publications and press releases. It is also important for the information to be expressed in nontechnical terms capable of being understood by the non-specialist user. Such information should either accompany the data disseminated or be provided in clearly linked technical notes.

- Where possible, quantitative measures of non-sampling error should be provided. However, because of the difficulty in quantifying some non-sampling errors, agencies will need to disseminate a mixture of quantitative and qualitative information that enables a non-technical user to clearly understand the strengths and limitations of the data. In particular, information on nonsampling errors should clearly convey to the user the fact that such errors, either individually or in total, may have a greater impact on the reliability of the data than sampling error and that the "ready" availability of quantitative measures of sampling error is not necessarily an indication of their relative significance.
- With respect to precisely what information on non-sampling errors that should be reported, the ideal recommendation is for national agencies and international organisations to disseminate information on all of the non-observation and observation errors summarized above. The secondbest option is the adoption of a more pragmatic approach which entails national agencies using their professional judgement and more detailed knowledge about the data to identify a sub-set of key non-sampling errors that have a significant impact on the reliability of the data in question. The important thing is for these agencies to develop a culture of critical appraisal of their statistical output and for key strengths and weaknesses to be documented and disseminated.

7.4.5 Recommended practices for rebasing

It is recommended that rebasing be undertaken every five years and within three years from the end of the base year. Unless the year was "unusual" it is also recommended that the base year selected be one ending with a "0" or "5". In order to provide sufficient transparency to users with regards to a rebase it is necessary to ensure that the following metadata accompanies any rebased data, either directly or through the provision of: appropriate references or links:

- the methodological approach adopted for the rebase, in particular, the processes actually undertaken during the rebase, e.g. simple referencing, introduction of new weights, etc.;
- the link year;
- the classification level at which index numbers are rebased and disseminated;
- the rounding policy followed in the rebasing, even though rounding should only be carried out at the very last stage for presentation purposes;
- a transition table from the old to the new classification system, if this is introduced;
- the description of any new weighting system and its impact on the aggregation of lower level indices;
- when the direct approach is adopted, a note of caution is useful to alert users that any aggregation of rebased indices needs the updating of the weights of the previous bases.

7.5.4 Recommended practices for citation of datasets

If citation of datasets is to be taken seriously, a concerted effort must be made by national agencies and international organisations to:

- Formulate and then place their data citation policy in an obvious position on websites, including the policy for the citation of data disseminated via electronic datasets. Furthermore, this policy should be accompanied by detailed sample citations to be included in specific web pages for users to copy as required. This makes it easy for users to include the correct information in the citation. The following example of this practice is from Statistics Canada's Census web module Community Profiles where the following citation instruction is provided at the bottom of the webpage in printer friendly format:

How to cite: Statistics Canada. 2002. 2001 Community Profiles Released June 27, 2002. Last modified: 2005-11-30. Statistics Canada Catalogue no. 93F0053XIE.

- Secondly, encourage a culture of data citation both inside and outside the organisation wherever data is being used. This awareness can be raised by contacting all known users of an organisation's data, all editors of publications known to use an organisation's data, etc., requesting that they follow the citation policy for the organisation in future publications.

A simple but effective citation style for datasets would be to include the following elements:

- unambiguous name of the dataset;
- author of the dataset;
- agency (or part of the agency) responsible for the dataset;
- date of the dataset (or version number);
- contact details for queries;
- address of the archive or other place of storage or system for accessing data;
- publisher (if this is different from the author, though for many agencies' publications the author and publisher are the same);
- if appropriate, the paragraph, table or page number.

This citation style should be followed for any data that is published internally or externally as well as for the documentation of any datasets that are created or modified. The actual ordering of the elements outlined above, punctuation, use of italics, etc., is a matter of individual (or organisational) choice.

Citation of text

The main recommendation for text citation entails the systematic use in all metadata of one of the widely accepted bibliographic reference styles listed in Section 7.5.2. The two commonly used systems for presenting references in text for a bibliography are the Harvard system and the Numeric system. It is beyond the scope of the current Handbook to outline these systems in any detail beyond outlining a number of specific areas in metadata presentation where such systems should be used. These include the provision of:

- References or source for concept or variable definitions used in all published output, e.g. definitions appearing in explanatory notes, glossaries, etc. At the moment it is almost impossible to identify the primary source of concept and variable definitions published by both national agencies and international organisations. In particular, it is seldom possible for the user to identify whether or not a specific definition: has been taken directly from existing international statistical standards; is a modified version adapted for a specific use (say at the national level); or an entirely new definition.

- Sufficient reference (citation) information to enable the user to readily identify the availability of more detailed information on definitions and concepts, collection methodology, etc. An example of a clear statement of citation policy at the national level is provided in Statistics Canada's publication, How to Cite Statistics Canada Products (Statistics Canada 2006a), which provides examples of recommended citation practices for a wide range of statistics products: publications; data products; census products; microdata products; maps and geospatial products; and E-STAT products.

7.6.2 Recommended practices for the presentation and reporting of administrative data

Because statistics derived from administrative sources will be based on data that were not originally compiled or produced for statistical purposes and frequently by other non-statistical agencies, Statistics Canada in their policy guidelines for informing users on data quality and methodology (Statistics Canada 2000, p. 12) states that it is particularly important for such data to be methodologically transparent to users and stressed the need for such data to accompanied by the following types of metadata:

- the name of the source agency for the administrative data. If more than one agency or ministry provides the services and collates data on these (e.g. health or education services provided by several agencies in some countries) specific information should be provided as to whether or not the data are from all agencies, or only from the main agency or ministry;
- a precise description of the purposes for which the statistical data were originally compiled and collected by the administrative agency;
- an outline of the strengths and weaknesses of the data in terms of the statistical application of the data. Particular attention should be given to the impact of issues relating to coverage and possible coverage bias, differences in concepts from international statistical guidelines and recommendations, in particular, the use of non-standard classifications and the use of unit concepts that differ from statistical units concepts;
- a description of processing or transformation (if any) undertaken by the statistical agency following receipt of the administrative data. Such processing may attempt to reduce or minimise inherent weaknesses in the original data;
- descriptions of the reliability of the data, including adherence to international norms and standards and caveats / limitations on the statistical use(s) of the data, e.g. for social indicator generation.

Corporate glossaries should contain the following in order to facilitate their interoperability: concept label; definition; detailed source information about where the definition was derived; related terms; and context field providing additional information or highlighting how a definition is used within one statistical domain or geographical context. Finally, the glossary should also be translated into various languages used by the institution in question.

Glossary

This Glossary intends to support the reader in understanding the terminology related to seasonal adjustment. The definitions are mainly drawn from international recommendations, most often they are quoted word for word and reference is provided. In some cases, they have been modified to suit the context of this *Guide*. These definitions mostly include a practical explanation of a term rather than an exact scientific definition.

Additive outlier	An additive outlier is an impulse outlier which affects only the value of one observation. Random effects, strikes or bad weather might cause this kind of outliers. A pre-announced price rising could cause an additive outlier by increasing the sales dramatically before the price change.
Additive decomposition	The additive decomposition means that the time series adds up as a sum of its independent components: trend-cycle + seasonal + irregular(/transitory). For additive time series, the magnitude of seasonal or irregular variations doesn't change as the level of the trend-cycle changes. Usually, any series with zero or negative values are treated in seasonal adjustment as additive series.
Advance release calendar	An advance release calendar provides a general statement on the schedule of release of data, which is publicly disseminated so as to provide prior notice of the precise release dates on which a national statistical agency, other national agency, or international organization undertakes to release specified statistical information to the public. (OECD, 2007.)
Autocorrelation	Autocorrelation is the correlation within a time series, or a signal, with its own past and future values. In other words, it refers to linear dependence between the values for different periods of a stationary variable. It is also sometimes called "lagged correlation" or "serial correlation".
Auto-regressive spectrum	An auto-regressive spectrum is a spectral plot used to alert the user about the presence of remaining seasonal and trading day effects (Grudkowska, 2011). In Demetra+, seasonal frequencies are marked as grey vertical lines in the auto-regressive spectrum, while the purple lines correspond to trading day frequencies. Peaks at the seasonal or trading day frequencies indicate the presence of seasonality or trading day effects.
ARIMA model	ARIMA models (p,d,q) are used for modelling and forecasting time series data. ARIMA stands for an Auto-Regressive Integrated Moving Average. The ARIMA model includes three types of parameters: the autoregressive parameters (p), the number of differencing passes (d) and moving average parameters (q). (p,d,q) is an abbreviation of the order of a model. A seasonal series usually has two sets of these parameter types: a regular component defined by (p,d,q), and a seasonal component (P,D,Q). For example, the Airline model (0,1,1)(0,1,1) includes no autoregressive parameters, but it includes one regular and one seasonal moving average parameter which were calculated after the series was differenced once for the regular component, and once it was seasonally differenced. Additionally, Reg-ARIMA would refer to the regression part which enables the use of regression variables.
Base period	The base period, usually a year, is generally understood to be the period against which other periods are compared and whose values provide the weights for an index. However, the concept is not precise, and three types of base periods may be distinguished: the quantity reference period, the weight reference period and the index reference period. See these terms below. (UNSD, 2010.)

Box-Pierce on residuals This test examines evidence of autocorrelation in the residuals. A rejected null hypothesis signifies evidence of autocorrelation. As stated before, this indicates a remaining linear, unwanted structure in the series, i.e. outstanding information.

Calendar effects Calendar effects refer to any economic effect related to the calendar, and are removed before seasonal adjustment and decomposition of a series. Calendar effects typically include: the different number of working days in a specific month or period, the composition of working days, the leap year effect and the moving holidays such as some national holidays (e.g. Easter and Ramadan) (Eurostat, 2009). Usually, calendar effects include both working day effects and trading day effects.

Calendar adjustment Calendar adjustment, within the pre-treatment of seasonal adjustment, deals with the non-seasonal part of calendar effects. Part of calendar effects are seasonal, i.e. the repeating length of months and the Easter falling most often in April. These seasonal effects are removed by standard seasonal adjustment, not by calendar adjustment. (Eurostat, 2009.) Usually, calendar adjustment includes both working day adjustment and trading day adjustment.

Canonical decomposition Canonical decomposition maximises the variance of the irregular component, which also maximises the smoothness of the seasonal and trend-cycle component so that no additive white noise can be removed from them. Each component of a series should follow an ARIMA model, except for the irregular component assumed to be white noise.

Chain-linking Joining together two indices that overlap in one period by rescaling one of them to make its value equal to that of the other in the same period, thus combining them into single time series. More complex methods may be used to link together indices that overlap by more than one period. (ILO, 2004.) Unlike the fixed weight approach, the chain approach doesn't re-calculate the entire historical series whenever the weights are updated but rather links or splices together the two index series to produce a coherent time series (UNSD, 2010).

Current adjustment Current adjustment is one of the alternative refreshment strategies in seasonal adjustment. This means that the model, filters, outliers and regression parameters are re-identified, and the respective parameters and factors re-estimated at review periods that have been set in advance. This approach makes use of the forecasted seasonal and calendar factors until the next review period. (Eurostat, 2009.) The reviews usually take place once or twice a year. In other words, current adjustment implies that the seasonal and calendar factors applied to new raw data in-between the review periods are fixed. Due to the forecasts, this strategy can lead to a lack of precision in the estimation of the latest adjusted figures.

Controlled current adjustment Controlled current adjustment is one of the alternative refreshment strategies in seasonal adjustment. It is one of the balanced alternatives between current and concurrent adjustment. In this option, forecasted seasonal and calendar factors derived from a current adjustment are used to seasonally adjust the new or revised raw data. However, an internal check is performed against the results of the "partial concurrent adjustment" which is preferred if a perceptible difference exists. This means that each series needs to be seasonally adjusted twice. The approach is only practicable for a limited number of important series. (Eurostat, 2009.)

Concurrent adjustment Concurrent adjustment is one of the alternative refreshment strategies in seasonal adjustment. This means that the model, filters, outliers and regression parameters are re-identified with the respective parameters and factors every time new or revised data become available (Eurostat, 2009). This adjustment strategy can lead to instability of the seasonal pattern, although it produces accurate results relying on the available data for each period. Often other balanced alternatives between this approach and current adjustment are used.

Cross-correlation Cross-correlation is a standard method of estimating the degree to which two series are correlated. In seasonal adjustment, the theoretical components of a time series are assumed to be uncorrelated. Each component is expected to follow an ARIMA model, except for the irregular component which consists of white noise. The estimators of the components are correlated, but the correlation is usually small. Thus, the quality diagnostics include tests on the cross-correlation of the estimators and actual estimates of components.

Cumulative indices Cumulative indices are a measure of summarizing the development of the current year compared with the corresponding period of the previous year. A cumulative figure may provide data for example for a period from January to May. In the next release, data would cover the period from January to June. Thus, the length of the reference period changes with each release of data, and the user cannot match revisions to the correct month. Cumulative indices are useful as additional information, not as the main release format. See year-to-date growth rates.

Decomposition Decomposition refers to the act of splitting a time series into its constituent parts by using statistical methods (OECD, 2007). Typically seasonal adjustment separates the trend-cycle component, the seasonal component and the irregular component from the original series. Sometimes a transitory component may also be identified by SEATS. Once the components have been estimated, the irregular and transitory components can be put together into an irregular/transitory component. The components usually form the original series either by adding up, i.e. by an additive decomposition, or it can be formed by multiplying the components, i.e. by a multiplicative decomposition.

Diagnostics Diagnostics are used to ensure that seasonally adjusted data are of good quality. Among others, the absence of residual seasonal and/or calendar effects as well as the stability of the seasonally adjusted pattern has to be carefully assessed. The validation of seasonally adjusted data can be performed by means of several graphical, descriptive, non-parametric and parametric criteria included in the output of the seasonal adjustment program. If possible this could be complemented with graphical diagnostics. (Eurostat, 2009.)

Differencing A time series which is not stationary with respect to the mean can be made stationary by differencing. Regular and seasonal differencing are often used to transform the series, removing trend and seasonal effects, in order to achieve a clearer view of the underlying behaviour of the series. See stationary below.

Fisher index formula Fisher index is defined as the geometric average of the Laspeyres index and the Paasche index. It is a symmetric and superlative index. (ILO, 2004.) See Laspeyres and Paasche index below.

Fixed weight indices Fixed weight indices have a weight structure selected at a particular point in time. The weights represent the relative importance of different products or activities at the weight reference period. Weights are used to compute indices over an extended period. This index compares the values always to the fixed base period. The weights are typically updated every five years (sometimes once a year), and at the same time ideally the entire time series is re-calculated based on the new weights. (UNSD, 2010.)

Flow series Flow series measure the activity over a given period (UNSD, 2010). An example of a flow series could be the retail trade statistics: the daily sales are summed up to give a total value for sales for a given month.

Friedman test The seasonal component includes the intra-year variation that is repeated constantly (stable seasonality) or that evolves from year to year (moving seasonality). A Friedman test determines if stable seasonality is present in a series. If the p-value is lower than 0.1% the null hypothesis of no stable seasonality is rejected; Otherwise the series is considered non-seasonal. (Grudkowska, 2011.) The Friedman test is a non-parametric statistical test developed by the U.S. economist Milton Friedman.

Index The name "index" comes originally from Latin and means a pointer (UNSD, 2010). An index is a ratio that indicates the increase or decrease of a magnitude (Allen, 1975). The index form is used not only for intertemporal comparisons but for comparisons between countries (Balk, 2008).

Index reference period The period for which the value of the index is set equal to 100 (UNSD, 2010).

Irregular component The irregular component captures the remaining short-term fluctuations in the series which are neither systematic nor predictable. It is assumed to include only white noise. When it contains short-term fluctuations that are not white noise, nor can they be assigned to the trend-cycle or seasonal component, SEATS captures these stationary fluctuations into a transitory component. The transitory component, if any, displays highly short-term behaviour, and will be added back to the irregular component, once all components of the series have been estimated. The irregular component is derived after the seasonal (including the calendar effect) and trend-cycle components have been removed from the original data (OECD, 2007). It can contain both random effects, i.e. white noise and some errors which are not necessarily random.

Kurtosis Kurtosis is a statistical measure which describes the distribution of the observed data around the mean. A rejected null hypothesis signifies that there's kurtosis in the residuals. Kurtosis is a measure of how peaked or flat a distribution is relative to the normal distribution. The higher the value the more peaked the data are. The term was introduced by Karl Pearson in 1906[18].

Kruskall-Wallis test A Kruskall-Wallis test is a non-parametric test on stable seasonality that compares samples of two or more groups to see whether the samples originate from the same distribution. The null hypothesis states that all months or quarters have the same mean. (Grudkowska, 2011.)

Laspeyres index In the Laspeyres index formula the price of a commodity at a given month is multiplied by the quantity at a base period. The value of the commodity basket at the current period, based on the quantities of the base period, is divided by the value of the commodity basket at the base period. Also called a base weighted index (ILO, 2004).

Level shift A level shift is a type of outlier that refers to a change into higher or lower values, and the level of the time series undergoes a more permanent change. Level shifts may happen due to a changed economic behaviour, new legislation, changed social traditions or changed definitions or classifications. For example, when the salaries increase for some profession, the level of that time series becomes higher permanently, but the seasonal pattern does not change.

[18] http://stats.oecd.org/glossary

Ljung-Box on residuals — Ljung-Box on residuals is a statistical test for identifying residual autocorrelation. If the residuals are uncorrelated and normally distributed, they also are independent. A rejected null hypothesis signifies evidence of autocorrelation in the residuals. This indicates a remaining linear, unwanted structure in the series, i.e. outstanding information instead of residual noise.

Log-transformation — The term log-transformation refers to computing a logarithm of each value in the series. It can simplify the statistical models. Especially the series with multiplicative decomposition are easier to work with after a log-transformation. The logarithm tends to bring together the larger values and stretch out the smaller values, which may reduce problems related to skewness, outliers or variation. This makes the series more stationary. Seasonal adjustment software usually provides an automatic test for log-transformation. See skewness and stationarity below.

Mean correction of residuals — If a seasonal model is well identifiable the residuals will follow normal distribution, and their mean should be zero. If not, a mean correction is performed, e.g. by TRAMO/SEATS or X-12-ARIMA, to bring the mean of residuals to zero.

Metadata — Metadata provide information on data and about processes of producing and using data. Metadata describe statistical data and - to some extent - processes and tools involved in the production and usage of statistical data. (UNECE, 1995.)

Moving average — A moving average is a method for smoothing time series by averaging (with or without weights) a fixed number of consecutive terms. The averaging “moves” over time, in that each data point of the series is sequentially included in the averaging, while the oldest data point in the span of the average is removed. In general, the longer the span of the average, the smoother the resulting series is likely to be. (OECD, 2007.) Moving averages are sometimes referred to as sliding spans. The moving average means, for instance, that one calculates for December 2010 an average of observations from January 2010 to December 2010, and for January 2011 an average of February 2010 – January 2011.

Moving holidays — Moving holidays are holidays which occur each year, but where the exact timing shifts. Examples of moving holidays include Easter, Chinese New Year and Ramadan. (OECD, 2007.)

Multiplicative decomposition — Multiplicative decomposition means that the original time series is the product of its components: trend-cycle x seasonal x irregular(/transitory). The multiplicative decomposition implies that as the trend of the series increases, the magnitude of the seasonal spikes also increases. Most economic time series exhibit a multiplicative relation between the components.

Normality of residuals — Normality test is a standard statistical test of the assumption that the model residuals are normally distributed. This assumption is needed for estimation of the model, for the adequacy of several of the diagnostics and for valid inference. Rejection of the hypothesis indicates that non-linear effects are present. Those affecting asymmetry (skewness) are more damaging than those affecting the thickness of the tails (kurtosis).

Null hypothesis — Scientific practice formulates and tests hypotheses, assumptions that can be proven correct or false using a test of the observed data. The null hypothesis typically corresponds to a general or default position, for example, that the components of a time series are uncorrelated.

Original series — Original series contains all characteristics of a series, without any adjustments or reduction. In the context of this publication, original series refers to the raw data that have not been transferred and may be expressed as values or indices. Both random and systematic fluctuations influence the original series. It may contain seasonal effects and effects related to the calendar.

Outlier — Outliers are abnormal values of the series. They can manifest themselves in a number of ways, the most important being additive or impulse outliers (abnormal values in isolated points of the series), transitory changes (series of outliers with transitory effects on the level of the series), and level shifts (series of outliers with a constant and permanent effect on the level of the series). (Eurostat, 2009.)

Paasche index — In the Paasche index formula the price of a commodity at a given month is multiplied by the quantity at the current period. The value of the commodity basket at the time of calculation is divided by the sum of the products of prices at base period and quantities at the current period.

Partial concurrent adjustment — Partial concurrent adjustment is one of the alternative refreshment strategies in seasonal adjustment. It is one of the balanced alternatives between the current and concurrent approaches. Partial concurrent adjustment means that the model, filters, outliers and calendar regression variables are re-identified, usually, once a year, but that the respective parameters and factors are updated every time new or revised observations become available (Eurostat, 2009).

Period-to-period growth rates — Period-to-period growth rates are rates of change expressed with respect to the previous period, for example, from April to May. They may be referred to as month-on-previous-month growth rates, month-over-month growth rates, 1-month rate of change, or rate of change on the previous month. Such rates are expressed as (M_t/M_t-1) -1 or (Q_t/Q_t-1) -1. Mt denotes the value of a monthly time series in month t and Q_t the value of a quarterly time series in quarter t. (OECD, 2007.)

Periodogram — A periodogram is a spectral plot used to alert the user about the presence of remaining seasonal and trading day effects (Grudkowska, 2011). In Demetra+, the seasonal frequencies are marked as grey vertical lines in the periodogram, while the purple lines correspond to trading day frequencies. Peaks at the seasonal or trading day frequencies indicate the presence of seasonality or trading day effects.

Pre-treatment — The main objective of pre-treatment of a series is to ensure a reliable estimation of the seasonal component. This is done in particular by detecting and correcting the series for data and/or components, sometimes called "non-linearities" that could hamper the estimation of the seasonality. (Eurostat, 2009.) These include outliers and the non-seasonal part of calendar effects. In addition, seasonal adjustment relies on the stationarity of the series, achieving this may require transformation of the series (Eurostat, 2009). Pre-treatment usually includes a test for log-transformation. After pre-treatment of the series, SEATS or X-11 separates the linearised series into trend-cycle, seasonal and irregular component.

Quantity (price) reference period — The quantity reference period is the period, whose volumes appear in the denominators of the volume relatives used to calculate the index (UNSD, 2010). For price indices, this is called the price reference period (ILO, 2004).

Rebasing There's some ambiguity in the concept of the base year. Rebasing may mean
• Changing the weights in an index,
• Changing the price or quantity reference period of an index number series, or
• Changing the index reference period of an index number series.
The weights, the quantity or price reference period, and the index reference period may be changed at the same time, but not necessarily so. (ILO, 2004)

Refreshment strategy In this context, refreshment strategy refers to the choice of method for updating the seasonally adjusted series when new or revised observations accumulate. The choice of refreshment strategy affects the frequency of revisions and the accuracy of seasonally adjusted data. The most common alternatives are called current adjustment (fixed), controlled current adjustment, partial concurrent adjustment and concurrent adjustment, where everything is revised during each update just as in the first seasonal adjustment of any time series.

Regression variable A regression variable explains another variable, for instance in seasonal adjustment an outlier or a calendar effect. You can apply regression variables for modelling the impact of moving holidays, working and trading days. TRAMO/SEATS and X-12-ARIMA remove the effect of the regression variables to estimate the seasonal pattern, but some of these effects are returned to the seasonally adjusted series, for example, outliers. You should apply a regression variable only if the series is influenced by the effect in question, and if the statistical tests prove it.

Residual seasonality Residual seasonality test is one of the seasonality tests included in Demetra+ for assessing the validity of results. Residual seasonality is an undesired feature in the seasonally adjusted series and in the residuals. The presence of residual seasonality is a concrete risk which could negatively affect the interpretation of seasonally adjusted data (Eurostat, 2009).

Revisions Data revisions are defined broadly as any change in a value of a statistic released to the public by an official national statistical agency. (OECD, 2007). The revisions are caused by changes in the unadjusted original data and the Reg-ARIMA model, and also by the two-sided filter and the use of forecasts in seasonal adjustment. The unadjusted data revise due to the corrections and accumulation of raw data, and the model because of the revisions incorporated by new observations. The two-sided filter means that the estimator of the seasonal component depends on the observations prior to and after a certain period. For the recent period, it implies dependence on observations that need to be forecasted. A revision incurs once the estimators based on forecasts are revised with new observations.

Seasonal adjustment Seasonal adjustment is the process of estimating and then removing from a time series influences that are systematic and calendar related. Observed data need to be seasonally adjusted, as seasonal effects can conceal both the true underlying movement in the series as well as certain non-seasonal characteristics which may be of interest to analysts.[19]

Seasonal component Seasonal component refers to those fluctuations observed during the year which repeat on a fairly regular basis from one year to another. The intra-year fluctuations are more or less stable year after year with respect to timing, direction and magnitude (OECD, 2007). The seasonal component depicts

[19] www.abs.gov.au/websitedbs/D3310114.nsf/Home/Methods,+Classifications,+Concepts+&+Standards

systematic, calendar-related movements in a time series. These regular movements don't give a clear indication of the underlying long-term development which is why they are removed during seasonal adjustment.

Seasonality The fluctuations in a time series are either repeatable or non-repeatable. By definition, seasonality is a pattern of a time series in which the data experience regular and predictable changes that repeat every calendar year. Seasonality refers to periodic fluctuations, e.g. the tendency of retail sales to peak during the Christmas season. Human activity includes rhythms, such as the 24-hour rotation of days and social habits, such as the alternation between night and day, one day of week of rest, holiday periods and consumption habits depending on the season. Seasonality reflects traditional behaviour associated with the calendar. It may also include the impact of business habits, such as quarterly provisional tax payments or periodic invoicing and administrative procedures, such as timing of tax returns.

Signal extraction In the context of time series, signal extraction refers to the estimation of the signal from the data. In SEATS signal extraction implies minimum mean square error estimation of the signal (seasonal and other components) contained in the model identified for the observed series. Sometimes also called smoothing of the data.

Skewness Skewness is a measure of the quality of the results of seasonal adjustment. It is a measure of how symmetrical a distribution is. A rejected null hypothesis would signify some evidence of skewness in the residuals. This means that the residuals would be asymmetrically distributed. A symmetrical distribution has a skewness factor of zero.

Sliding spans See moving average.

Stationarity Stationarity is a set of conditions that allow estimation of model parameters whose properties are standard. In other words, a stationary time series is easier to model and forecast. A stationary time series should have a constant mean, variance and autocorrelation through time. If differencing is required to achieve stationarity, the model is called an autoregressive integrated moving average or ARIMA. [20]

Stock series Stock series are measures of a variable taken at points in time (UNSD, 2010). The monthly labour force survey is a usual example of a stock series. It takes stock of whether a person was employed in the reference period or not. Often, stock series may be measured at the end of the reference period, for example, money supply data which can refer to an observation on the last working day of the reference period (OECD, 2007).

Time series A time series is a collection of observations for a variable over time. It is measured at regular time intervals, e.g. monthly or quarterly. For seasonal adjustment, a time series has to be measured at time intervals shorter than one year, because seasonal fluctuations are intra-annual, and tend to repeat one year after the other. In addition, the series has to be compiled for discrete periods of time, e.g. for every month or quarter.

Trading day effect Recurring day-of-week effects and leap year effects in monthly (or quarterly) economic time series are called trading day effects (Findley and Soukup, 2000). The effects of trading days are estimated by counting the proportion of them in

[20] http://www2.sas.com/proceedings/sugi30/192-30.pdf

	the month (or quarter) (Eurostat, 2009). Thus, trading day effect means that the level of the measured activity varies depending on the day of the week. For example, sales may be higher on Fridays than on Tuesdays.
Trading/working day adjustment	Trading/working day adjustment aims at obtaining a seasonally adjusted series whose values are independent of the length and the composition in days (number of Mondays, Tuesdays, etc/number of working days and weekend days) of the month (or quarter) (Eurostat, 2009). Most often, trading and working day adjustment are used as synonyms, but in some cases, working day adjustment focuses on the difference between working days and non-working days, whereas trading day adjustment refers to the effect of different days of the week.
Transitory component	An additional component that may be used by SEATS. It picks up highly transitory variation that is not white noise and should not be assigned to the seasonal or trend-cycle component. Its model is a low-order stationary ARMA model, with spectral peaks at non-seasonal and non-trend frequencies. Not all series need this additional component.
Transitory change	A transitory change is a type of outlier, namely a point jump followed by a smooth return to the original path. They remain visible in the seasonally adjusted series. Transitory changes may occur, for example, due to deviations from average monthly weather conditions. If in the winter the weather becomes suddenly colder it may lead to a peak in energy consumption. When the weather gradually returns to the regular level the consumption should settle back to normal.
Trend-cycle component	The trend-cycle component includes both very long-term and medium-term developments. Out of these two, the trend component depicts the long-term evolution of the series that can be observed over several decades, i.e. the structural variations. It reflects the underlying level of the series, and is typically the result of influences such as population growth, price inflation, technological development and general economic development. The cycle component, on the other hand, is the relatively smooth movement around the long-term trend. It is a rhythmic cycle caused by economic variation from expansion to recession.
Year-on-year growth rates	Year-on-year growth rates are rates of change expressed over the same period, month or quarter, of the previous year. They may be referred to as year-over-year growth rates, year-to-year growth rate, rate of change from the previous year, or 12-month rate of change. Such rates are expressed as $(M_t/M_t\text{-}12)\text{-}1$ or $(Q_t/Q_{t\text{-}}4)\text{-}1$. (OECD, 2007.)
Year-to-date growth rates	Year-to-date data are expressed in cumulative terms from the beginning of the year; sometimes referred to as cumulative data (OECD, 2007). Year-to-date growth rates are data expressed in cumulative terms from the beginning of the year and compared with the same period of the previous year. For example, they may compare the sum of values from January 2011 to April 2011, to the same period of 2010.
Weight reference period	The weight reference period is the period, usually a year, whose values serve as weights for the index. (UNSD, 2010)

References

Allen R.G. D. (1975). *Index Numbers in Economic Theory and Practice*, New York.

Balk B. M. (2008). *Price and Quantity Index Numbers: Models for Measuring Aggregate, Change and Difference*, ISBN 978-0-521-88907-0, New York.

Bell W. R. and Hillmer S. C. (1984). *Issues Involved with the Seasonal Adjustment of Economic Time Series,* Journal of Business & Economic Statistics Vol. 2, No. 4, American Statistical Association.

Box, G. E. P. and Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control* (revised edition), Holden Day, San Francisco.

Central Bank of the Republic of Turkey (2002), *Seasonal Adjustment in Economic Time Series*, Istanbul. [www.tcmb.gov.tr/yeni/evds/yayin/kitaplar/seasonality.doc]

ECB (2000), *Seasonal Adjustment of Monetary Aggregates and HICP for the Euro Area*, Frankfurt. [http://www.ecb.int/pub/pdf/other/sama0008en.pdf]

ECB (2003), *Seasonal Adjustment*, edited by Michele Manna and Romana Peronaci, Frankfurt. [www.ecb.int/pub/pdf/other/statseasonaladjustmenten.pdf]

Eurostat (2009). *ESS Guidelines on Seasonal Adjustment*, Luxembourg, European Communities. ISBN 978-92-79-12307-8 [epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-RA-09-006/EN/KS-RA-09-006-EN.PDF]

Eurostat (2006). *Methodology of Short-term Business Statistics, Interpretation and Guidelines*, European Communities, Luxembourg. [epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-BG-06-001/EN/KS-BG-06-001-EN.PDF]

Findley, D.F., Monsell, B.C., Bell, W.R., Otto, M.C., and Chen, B.C. (1998). *New Capabilities and Methods of the X-12-ARIMA Seasonal Adjustment Program*. Journal of Business and Economic Statistics, 16, pp. 127 – 177 [www.census.gov/ts/papers/jbes98.pdf]

Findley, D.F., Monsell, B.C., Schulman, H.B. and Pugh, M.G. (1990). *Sliding-Spans Diagnostics for Seasonal and Related Adjustments*, Journal of the American Statistical Association, 85, 410, 345-355.

Findley, D. F. and Soukup R. J. (2000), *Detection and Modeling of Trading Day Effects,* ICES Proceedings. [www.census.gov/ts/papers/ices00_td.pdf]

Gómez, V. and Maravall, A. (1996), *Programs TRAMO and SEATS, Instructions for the User*, Working Paper 9628, Bank of Spain, Madrid.

Gómez, V. and Maravall, A. (2001). *Seasonal Adjustment and Signal Extraction in Economic Time Series,* Ch.8 in Peña D., Tiao G.C. and Tsay, R.S. (eds.) A Course in Time Series Analysis, New York: J. Wiley and Sons.

Grudkowska S. (2011). *Demetra+ User Manual*, National Bank of Poland. [www.cros-portal.eu/sites/default/files/Demetra+%20User%20Manual.pdf]

ILO, IMF, OECD, Eurostat, UNECE and World Bank (2004). *Producer Price Index Manual: Theory and Practice*, International Monetary Fund, Washington DC. [www.imf.org/external/pubs/ft/ppi/2010/manual/ppi.pdf]

Maravall, A., Pérez D. (2011). Applying and interpreting model-based seasonal adjustment, The euro-area industrial production series, Banco de España Working Papers, number 1116, Madrid. [http://www.bde.es/webbde/SES/Secciones/Publicaciones/PublicacionesSeriadas/DocumentosTrabajo/11/Fich/dt1116e.pdf]

OECD (2007), *Data and Metadata Reporting and Presentation Handbook*, Paris. [www.oecd.org/dataoecd/46/17/37671574.pdf]

UNECE (1992). *The Fundamental Principles of Official Statistics*, Geneva. [www.unece.org/fileadmin/DAM/stats/documents/e/1992/32.e.pdf]

UNECE (1995). *Guidelines for the Modelling of Statistical Data and Metadata*, Geneva. [www.unece.org/fileadmin/DAM/stats/publications/metadatamodeling.pdf]

UNECE (2009). *Making Data Meaningful, Part 2, A Guide to Presenting Statistics*, Geneva. [www.unece.org/fileadmin/DAM/stats/documents/writing/MDM_Part2_English.pdf]

UNSD (2010). *International Recommendations for the Index of Industrial Production*, United Nations Statistics Division, New York. ISBN 978-92-1-161532-6 [unstats.un.org/unsd/statcom/doc10/BG-IndustrialStats.pdf]

PRACTICAL GUIDE TO SEASONAL ADJUSTEMENT WITH DEMETRA+

FROM SOURCE SERIES TO USER COMMUNICATION

To assess the state of the economy and to make informed decisions on economic policy, correct and timely information has to be available about short-term economic development. However, since many economic phenomena such as production, income and employment are influenced by seasonal factors, simply relying on the raw, unadjusted statistical series may not give an accurate picture.

Seasonal adjustment attempts to remove the seasonal pattern of time series so as to reflect the underlying development. It makes it easier to draw comparisons over time and to interpret the development in the series. It also allows time series with different seasonal patterns to be compared between different industries or countries.

This *Practical Guide* introduces the reader to seasonal adjustment, giving examples of using Demetra+. It suggests an overall process for performing the adjustment and explains the related concepts. It also brings together some international recommendations on producing high-quality time series, performing seasonal adjustment and disseminating the results.