## OCCASION

This publication has been made available to the public on the occasion of the 50th anniversary of the United Nations Industrial Development Organisation.

## DISCLAIMER

This document has been produced without formal United Nations editing. The designations employed and the presentation of the material in this document do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations Industrial Development Organization (UNIDO) concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries, or its economic system or degree of development. Designations such as "developed", "industrialized" and "developing" are intended for statistical convenience and do not necessarily express a judgment about the stage reached by a particular country or area in the development process. Mention of firm names or commercial products does not constitute an endorsement by UNIDO.

## FAIR USE POLICY

Any part of this publication may be quoted and referenced for educational and research purposes without additional permission from UNIDO. However, those who make use of quoting and referencing this publication are requested to follow the Fair Use Policy of giving due credit to UNIDO.

## CONTACT

Please contact publications@unido.org for further information concerning UNIDO publications.

For more information about UNIDO, please visit us at www.unido.org

UNITED NATIONS INDUSTRIAL DEVELOPMENT ORGANIZATION
Vienna International Centre, P.O. Box 300, 1400 Vienna, Austria
Tel: (+43-1) 26026-0 · www.unido.org · unido@unido.org

UNITED NATIONS
INDUSTRIAL DEVELOPMENT ORGANIZATION

WORKING PAPER 09/2014

# Seasonal adjustment of national indices of industrial production at international level

INCLUSIVE AND SUSTAINABLE INDUSTRIAL DEVELOPMENT

RESEARCH, STATISTICS AND INDUSTRIAL POLICY BRANCH
WORKING PAPER 09/2014

# Seasonal adjustment of national indices of industrial production at international level

Shyam Upadhyaya

Shohreh Mirzaei Yeganeh
UNIDO
Statistics

UNITED NATIONS INDUSTRIAL DEVELOPMENT ORGANIZATION
Vienna, 2015

The designations employed, descriptions and classifications of countries, and the presentation of the material in this report do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations Industrial Development Organization (UNIDO) concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries, or its economic system or degree of development. The views expressed in this paper do not necessarily reflect the views of the Secretariat of the UNIDO. The responsibility for opinions expressed rests solely with the authors, and publication does not constitute an endorsement by UNIDO. Although great care has been taken to maintain the accuracy of information herein, neither UNIDO nor its Member States assume any responsibility for consequences which may arise from the use of the material. Terms such as “developed”, “industrialized” and “developing” are intended for statistical convenience and do not necessarily express a judgment. Any indication of, or reference to, a country, institution or other legal entity does not constitute an endorsement. Information contained herein may be freely quoted or reprinted but acknowledgement is requested. This report has been produced without formal United Nations editing.

This document reflects work in progress. Its distribution is limited for the purposes of eliciting comments and reviews only.

## Table of Contents

## 1. Background

Sub-annual macroeconomic statistics today are a key tool for economic policymaking, business cycle analysis and forecasting. However, these statistics are often influenced by seasonal fluctuations as well as calendar and trending day effects, which can cover relevant short- and long-term movements of the time series and obscure a clear understanding of economic phenomena.

Climatic and natural factors may influence consumption and production behaviour. Time series may also be affected by moving holidays such as Easter or Ramadan, and variations in fiscal years in different countries may bring seasonality into the time series. The calendar representing the number of working days in a month also causes movements in time series as the number of trading days differs from one month to another.

Hence, statistical indicators related to growth measurements are often characterized by significant seasonal variations and differences in the number of working days over time periods under comparison. The Quarterly Index of Industrial Production (QIIP) is one of the most important industrial short-term indicators measuring the rise and fall of industrial output volume with a special focus on detecting the business cycle’s turning points as early as possible. This allows planners, decision makers and the business community at large to detect signs of change in the economic progress in order to take appropriate and timely policy measures.

UNIDO collects high frequency data on IIP from National Statistical Offices (NSOs) and UNSD’s monthly bulletin. These data are exclusively obtained from secondary sources and compiled and disseminated by the NSOs. During the compilation process, all monthly data are converted into quarterly data by averaging the monthly values.

UNIDO Statistics has been publishing a series of quarterly reports on the world manufacturing growth rate based on quarterly compiled index numbers since the first quarter of 2011. The report’s main objective is to provide an overview of the current growth trends of world manufacturing production by country group and by major industries. Earlier reports included index figures for certain countries that were not seasonally adjusted or for which information was not available if seasonal adjustments were made at the national level.

For this reason, the results were presented in previous reports as a manufacturing growth percentage by linking the indices of the reference period to the previous periods selected for comparison:

1. as compared to the previous quarter
2. as compared to the same period of the previous year.

Both comparisons have their own use and advantages. The first set of growth indices represents more recent growth trends by comparing each quarter with the previous one. As national data are not all seasonally adjusted and considering the seasonal variation, the annual comparison (growth year compared to the same period of the previous year) provides more precise estimates on growth figures.

To obtain a meaningful comparison of growth rates for different periods within a given year and to better understand and interpret underlying trends UNIDO has been publishing growth figures based on seasonally adjusted index numbers. The index numbers are seasonally adjusted using the TRAMO/ SEATS method1 in the Demetra+ software.

The purpose of seasonal adjustment is to filter out seasonal fluctuations and remove the systematic calendar-related effects within the movements of the IIP time series. The reasons for seasonal adjustment of the IIP for UNIDO statistics can be summarized as follows:

- To provide more reliable short-term forecasts of industrial activities;
- To compare industrial production growth rates from different countries;
- To be able to adequately compare growth rates with the previous quarter;
- To record the real movements and turning points in manufacturing production, which may otherwise be impossible or difficult to determine due to seasonal movements.

Calculating the growth rate of a given period and comparing it to the same period of the previous year would be an implicit solution to avoid seasonality, but it could not be replaced by seasonal adjustment, since it does not remove all seasonal effects (e.g. moving holidays). Seasonally adjusted data provide more readily interpretable measures of changes that occur in a given period and reflect real economic movements without the misleading seasonal changes.

## 2. Basic concepts

This chapter reviews some basic times series concepts that are important for describing seasonal adjustment. In particular, we introduce the concept of time series components and seasonal adjustment methods.

## 2.1 Time series

A time series is a set of statistics, usually collected at regular intervals. Time series data occur naturally in many application areas such as monthly data for unemployment, the quarterly index of industrial production, the number of orders received/cancelled, etc.

Usually, the intent is to describe and summarize the data, apply a model and establish whether there is a discernible pattern among the values collected, with the intention of short-term forecasting (to be used as the basis of policymaking). We find:

$$\mathrm{Yt} \;=\; \mathrm{IIP\ at\ time\ t}$$

Time series data should be studied differently due to the dependency and correlation between the data.

The majority of statistical methods are based on the assumption of independence, which does not hold for time series data. An economic time series indicates the direction and turning points of economic activities. In addition, we can observe whether the series is increasing/ decreasing more slowly than before. As mentioned earlier, seasonal fluctuations can occasionally make it difficult or impossible to detect economic features in time series.

## 2.2 Components of time series

One way to describe a time series is to decompose the series into its components:

- Trend (Tt) – long-term movements at the level of the series
- Seasonal effects (St) – cyclical fluctuations related to the calendar (including moving holidays and working day effects)
- Cycles (Ct) – cyclical fluctuations longer than a year (such as business cycles)
- Irregular (It) – other random or short-term unpredictable fluctuations.

The idea behind seasonal adjustment is to create separate models for these components and then combine them additively:

$$Y_t \;=\; S_t + \,T_t \,+\, I_t \,+\, C_t$$

where the seasonal adjusted series is derived as:

$$SA\,_t \,=\, Y_t - S_t \;\; or\, SA\,_t \,=\, T_t + \, I_t$$

or multiplicatively:

$$Y_t \;=\; S_t - T_t - \, I_t$$

where

$$SA_t = \frac{Y_t}{S_t} = T_t \times I_t$$

The seasonal adjusted time series only contains irregular and trend components. The additive decomposition model assumes that the components of time series behave independently of each other. In other words, the amount of both the seasonal and irregular variations does not change as the level of the trend rises or falls. The observed time series in the additive model is considered to be the sum of the three independent components.

The multiplicative decomposition model is suitable for time series in which the magnitude of seasonal dips depends on changes in the components' trends, that is, for this series, the degree of seasonal variation increases as the level of the components' trends rises. In the multiplicative model, the original time series is expressed as the product of trends, seasonal and irregular components.

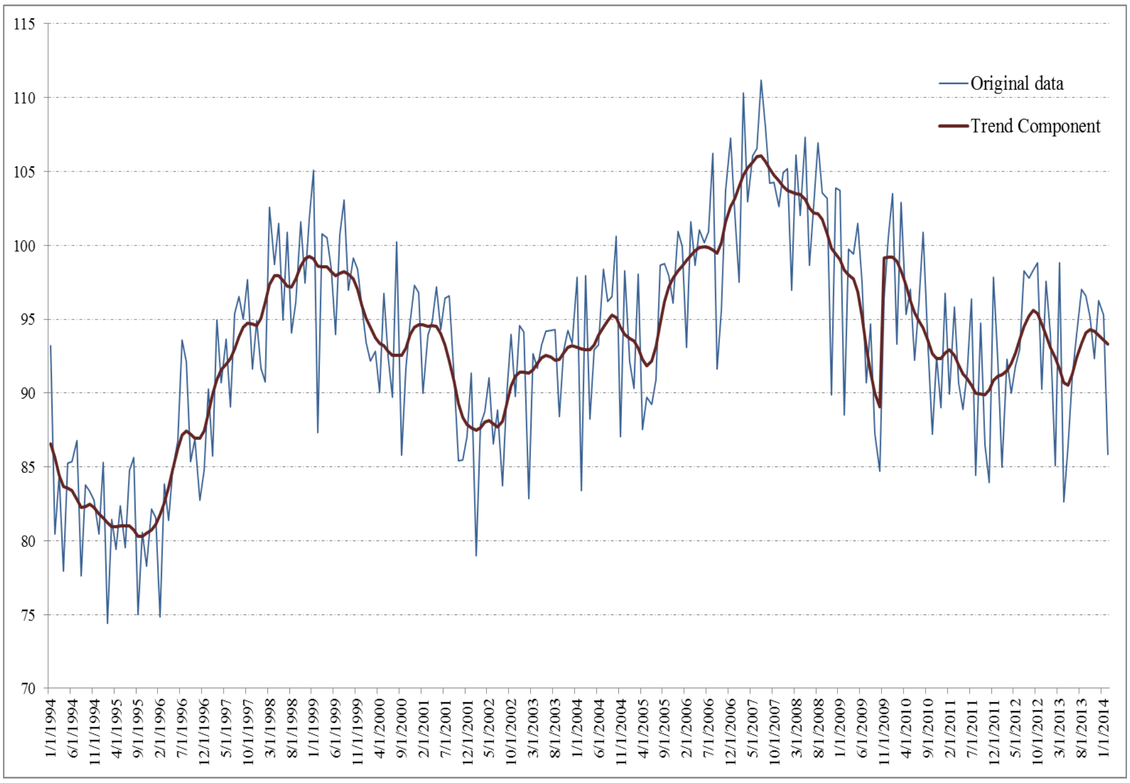
### *2.2.1 Seasonal effect*

Seasonality is a time series pattern in which data undergo regular and fairly predictable changes that recur every calendar year (periodic fluctuations, e.g. the Christmas effect, etc.). The presence of seasonality is quite common in economic time series. Seasonality may also entail calendar-related effects which do not occur regularly within a year due to variations in the calendar from year to year.

The possible causes of seasonal effects (S) can be classified into three groups: natural factors, administrative or legal rhythms and social/ cultural/ religious traditions (e.g. fixed holidays, timing of vacations). Seasonal effects in time series make it difficult to determine whether the changes are in fact an essential improvement or decline in the level of activity, or whether they are part of regular variations. Trading day and moving holiday effects are also two predictable and systematic calendar-related effects that are usually included in seasonal effects.

As mentioned earlier, seasonally adjusted time series only include trends and irregular components and exclude any seasonality. If no seasonality is present, the original data will be taken as seasonally adjusted data.

**Figure 1: IIP Example of trend component**



### *2.2.2 Trend*

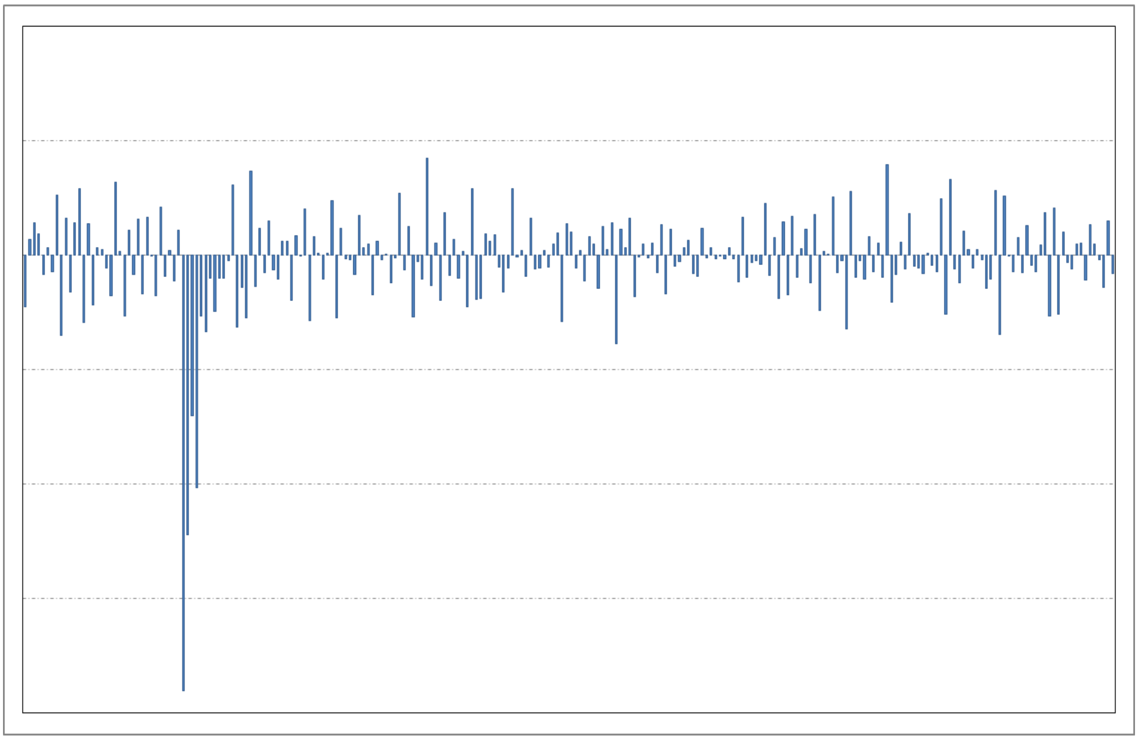
The trend component (plus cycles, also referred to as trend cycles) is a reasonably smooth time series which denotes the long-term movement and direction within a time series. The underlying direction of time series can be interpreted based on trends, with no influence of volatility which may distort and mislead the understanding of short-term movements. It could be regarded as the representation of the underlying level of the time series.

### *2.2.3 Irregular effect*

The irregular component (also referred to as the residual component) is the time series' unpredictable component, which will remain after excluding the seasonal (including calendar effect) and trend components from the original series. It attempts to capture the remaining short-term fluctuations which are not systematic or predictable. This effect is not predictable in terms of timing, impact and duration and includes errors of measurement and unusual events. The major causes of irregular effects are as follows: unseasonable (unexpected) natural disasters, strikes, sampling and non-sampling errors.

The seasonally adjusted series still contains an irregular component.

**Figure 2: IIP Example of irregular component**



### 2.2.4 *Other effects*

Trading day, moving holidays, the leap year effect and outliers are sometimes considered separate components of time series.

Trading (or working) day refers to the effect of the length of the month on the number of working days (see Table 1). The varying number of weekdays influences the economic time series each month. The number of Tuesdays in October 2013 was 5, for example, whereas it was only 4 the next month, November. The number of weekdays that occur five times within a month varies from month to month. Taking this factor into consideration, the trading day adjustment also assumes no economic activity on Sundays. For this reason, six regression variables for the remaining weekdays are used to adjust for this effect. Many real industrial activities are influenced by the trading day effect. Unlike the trading day, working day adjustment assumes no difference in economic activity between the working days, but between working and non-working days (Saturdays and Sundays). Thus, the varying number of such days must be considered. Most financial sector time series are influenced by the working day effect. Along with the abovementioned effects, the adjustment of calendar effects should include the leap year effect as well. The leap year effect may also impact the time series, since it adds an additional day every four years, which can be either a trading or a weekend day.

Moving holidays do not fall on the same time each year, but shift systematically. Since the number of holidays varies from one country to another, the prevailing national holidays should be considered for each country while seasonally adjusting the time series. Examples of moving holidays are Easter holidays, Ramadan and the Chinese New Year.

**Table 1: Calendar effect**

OCTOBER 2013

| M | T | W | T | F | S | S |
|---|---|---|---|---|---|---|
|  |  |  | 1 | 2 | 3 | 4 |
| 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| 12 | 13 | 14 | 15 | 16 | 17 | 18 |
| 19 | 20 | 21 | 22 | 23 | 24 | 25 |
| 26 | 27 | 28 | 29 | 30 | 31 |  |

NOVEMBER 2013

| M | T | W | T | F | S | S |
|---|---|---|---|---|---|---|
|  |  |  |  | 1 | 2 | 3 |
| 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 11 | 12 | 13 | 14 | 15 | 16 | 17 |
| 18 | 19 | 20 | 21 | 22 | 23 | 24 |
| 25 | 26 | 27 | 28 | 29 | 30 |  |

Trading days, moving holidays and leap years are calendar-related (i.e. predictable) effects. Therefore, these effects are usually considered to be included in seasonal effects.

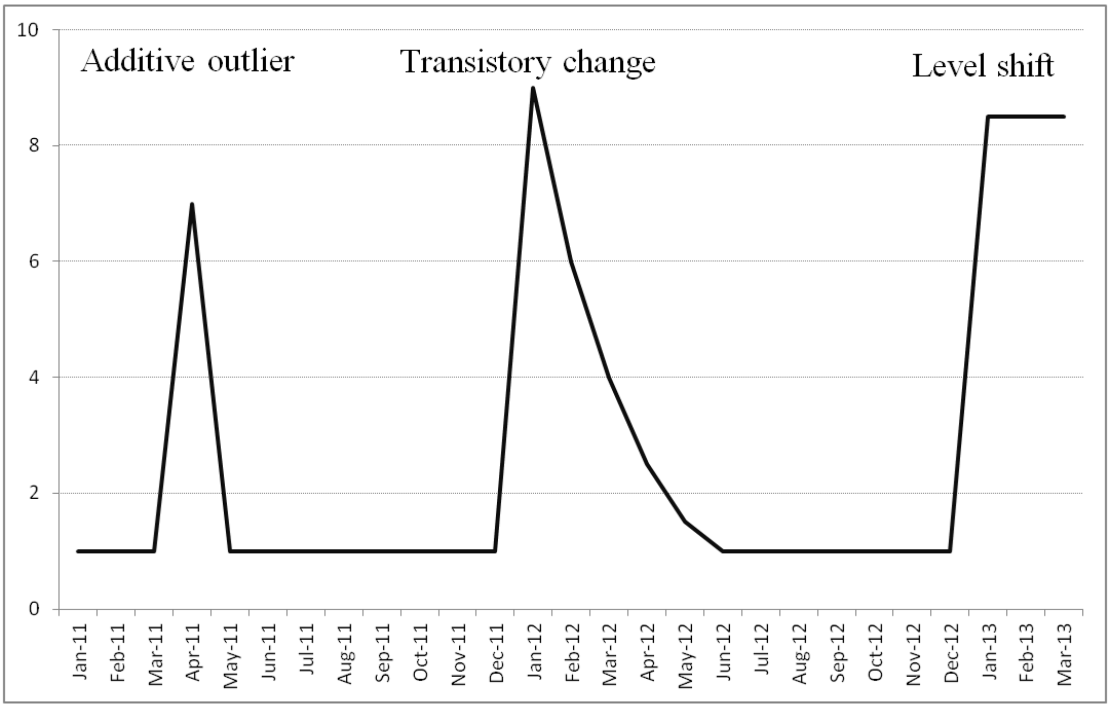
Outliers are extreme observations or records that deviate from the trend and fall outside the range expected from the typical pattern of time series. These abnormal values may occur as a result of new policies or new types of taxes, extreme natural events or closure of a significant manufacturer. The three most frequent outliers (see Figure 3) are additive outliers (AO), transitory (temporary) changes (TC) and level shift (LS).

**Additive outliers (AO):** This outlier type affects a single observation. After this interruption, the series returns to its normal path as if nothing happened. The AO may be caused by a random effect, a strike or a short-term shock in the system. For instance, a pre-announced price rise could be the cause of an AO by drastically increasing sales before the price change is even introduced. This outlier type must be identified because seasonal adjustment using TRAMO/SEATS is based on moving the average, and is therefore sensitive to extreme values.

**Transitory change (TC):** Transitory change outliers are spikes that take several periods to disappear exponentially. TC may occur due to deviations from average monthly weather conditions.

If the weather changes drastically, energy consumption may rise or fall. When the weather gradually stabilizes, energy consumption also returns to its usual level.

**Figure 3: Most frequent types of outliers**



**Level shift (LS):** Level shift refers to a more permanent change in the time series' level starting at a given time period. It may occur due to changes in concepts and definitions or compilation methods of the survey population; however, statisticians should try to avoid level shifts when changing methodology to preserve the comparability of the series. Level shifts may also occur as a result of changes in economic behaviour, in social traditions or in legislation. Level shifts change the level of the time series, but do not modify seasonal behaviour. These outliers must therefore be identified to avoid distorting the seasonal component.

Outliers contain much information about unusual events, so they are an important part of the data and will remain visible in seasonally adjusted time series.

## 2.3 Seasonal adjustment methods

In this section, a summary of currently used seasonal adjustment methods and software is presented. These methods can generally be classified into two main groups: model-based methods and filter-based methods. Both methods as well as other approaches will be explained below.

### *2.3.1 Model-based seasonal adjustment*

Model-based seasonal adjustment has increasingly been gaining attention. The model-based approach specifies explicit statistical models of the trend, seasonal and irregular components of the classical decomposition.

Examples of model-based methods include TRAMO/SEATS, X13-ARIMA/SEATS and STAMP.

TRAMO (**T**ime Series **R**egression with **A**RIMA Noise, **M**issing Observations and **O**utliers) and SEATS (**S**ignal **E**xtraction in **A**RIMA **T**ime **S**eries) were developed by Victor Gómez and Agustin Maravall at the Bank of Spain. The two programmes are being intensely used by data-producing and economic agencies, including Eurostat and the European Central Bank. TRAMO and SEATS provide a fully model-based method for forecasting and signal extraction in univariate time series. Due to their model-based features, they are powerful tools for a detailed analysis of time series.

### *2.3.2 Filter-based seasonal adjustment*

This method applies to a set of fixed filters (moving averages) to decompose the time series into a trend, seasonal and irregular component. Typically, symmetric linear filters are applied to the middle of the series and asymmetric ones to the end of the series. Filter-based methods of seasonal adjustment are often known as X11 style methods. They are based on the 'ratio to moving average' procedure. The procedure consists of the following steps:

1) Estimate the trend by a moving average
2) Remove the trend, leaving the seasonal and irregular components
3) Estimate the seasonal component using moving averages to smooth out the irregular components.

Examples of filter-based methods are X11-ARIMA and X12-ARIMA (uses regARIMA models for forecasts, backcasts and pre-adjustments).

## 3. Seasonal adjustment using Demetra+

When choosing a seasonal adjustment (SA) programme, statistical agencies had at least two different options in the past: X-12-ARIMA and TRAMO/ SEATS. Today, combined software packages exist, which merge the functionalities of X-12-ARIMA and TRAMO/SEATS. Demetra+ is a seasonal adjustment software that combines the two above mentioned methods.

The Demetra+ project is run by Eurostat. Unlike other software development efforts undertaken under an open source license, the Demetra+ project was not initiated by a community or a single developer, but started as an extension of the active role played by Eurostat in the promotion, development and maintenance of a statistical analysis software solution. The Seasonal Adjustment Steering Group (SASG), which consists of a Eurostat-ECB high level group of experts from NSIs and NCBs, has been promoting the development of freely available Demetra solutions for seasonal adjustment to be used within European statistical systems for several years. The SASG is also responsible for facilitating collaboration between different organizations interested in developing SA tools and supervises the entire project. Although the software itself is made available under an open source license, participation in its development is contingent upon a decision of the SASG. There are currently no plans for further open source development and testing. Development of the Demetra software has been outsourced to the Department of Statistics of the National Bank of Belgium (NBB), and, for the time being, NBB remains the sole developer of Demetra+. The User Testing Group has been set up primarily to supervise the implementation of the guidelines and user requirements. The User Testing Group is also responsible for issuing recommendations for new requirements and making decisions on the adoption or rejection of new requirements that are not in line with the project guidelines. The Demetra+ community builds on the OSOR environment in which the members of the User Testing Group can report and exchange experiences as well as communicate with the development team in NBB.

Demetra+ is a family of modules on seasonal adjustment, which are based on the two leading algorithms in that domain (TRAMO/SEATS and X-12-ARIMA). The two methods use a very similar approach in the first part of data processing, but differ completely in the decomposition part. Their comparison is often difficult, even for the modelling step. Moreover, their diagnostics focus on different aspects, and their outputs take completely different forms. One of the main features of Demetra+ is to normalize—to the extent possible—the different methods. It attempts to improve the two methods' comparability by using a common set of diagnostics and presentation tools. This fundamental decision implies that a number of procedures of both methods have been re-written in Demetra+. This could imply minor discrepancies in diagnostics or in peripheral information in comparison to the original programme, which should not, however, alter the general information provided by the algorithms. In any case, the original programme's main results (i.e. the seasonally adjusted series) should not be affected by this solution.

The literature does not propose any best method selection criteria for TRAMO/SEATS and X-12-ARIMA. Yet there are rigorous differences between the two methods. TRAMO/SEATS is a parametric method and has a model-based facility to extract the underlying components. The results of TRAMO/SEATS can be statistically explored. On the other hand, X-12-ARIMA is a non-parametric method and has an ad-hoc fixed filter facility to extract the underlying components. This means that the results of X-12-ARIMA cannot be statistically evaluated. However, there are also ad-hoc criteria to measure the reliability of seasonal adjustments of X-12-ARIMA.

UNIDO applies the TRAMO/SEATS method to produce seasonally adjusted IIP data using Demetra+, mainly due to the following reasons:

- Understandable statistical tests can be used and there is therefore no need to refer to ad hoc criteria (thus, subjectivity may be eliminated);
- Automatic model identification performance for the series is far superior to the alternative;
- The probability of finding spurious seasonality is lower than in the alternative;
- Demetra+ is a simple and easy to use interface and an excellent programme for non-professionals.

Statistical criteria can subsequently be used to measure the reliability of the process. The released Demetra+ 1.0.4 (17/12/2012) can be downloaded easily and is compatible with Microsoft Office 2010. The appropriate Microsoft Access DataBase Engine must be installed by the user to work with Excel workbooks. Detailed information on Demetra+ can be found in the user manual, available at the following link http://www.cros-portal.eu/sites/default/files//Demetra%2B%20User%20Manual%20November%202012.pdf.

The Demetra+ User Manual introduces the main features of the Demetra+ software. It presents an overview of the software's capabilities and functions. Moreover, a step-by-step description on how to carry out a standard analysis is provided. The manual describes the steps to follow and illustrates Demetra+'s user friendliness. The user is expected to have basic knowledge of seasonal adjustment and be familiar with the X-12-ARIMA and the TRAMO/SEATS methods. A brief outline of the X-12-ARIMA and the TRAMO/SEATS algorithms and concepts is also presented. References in the report provide the interested reader with additional sources on the seasonal adjustment concept.

## 3.1 Reg-ARIMA models in TRAMO

Let $z$ represent a time series observation

$$z = \left(z_{t_1}, z_{t_2}, \dots, z_{t_m}\right)$$

where $1 = t_1 < t_2 < \cdots < t_m = T$ are the time points (there may be some missing observations and the original observations may have been log transformed). The Reg-ARIMA model can be formulated as follows:

$$z_t = y'_t\beta + x_t \tag{1}$$

where $y_t$ is a matrix with $n$ regression variables and $\beta$ is the vector with the regression coefficients. The variable $x_t$ follows a (possibly non-stationary) ARIMA model. Hence, in (1), $y'_t\beta$ represents the deterministic component and $x_t$ the stochastic one.

If $B$ denotes the backward shift operator, such that $B^j z_t = z_{t-j}$, the ARIMA model for $x_t$ is of the type

$$v_t = \delta(B)x_t \tag{2}$$

$$\phi(B)[v_t - \mu_v] = \theta(B)a_t, \qquad a_t \sim Niid(0, V_a), \tag{3}$$

where $v_t$ is the stationary transformation of $x_t$, $\mu_v$ its mean, $\delta(B)$ is a non-stationary AR polynomial containing regular and seasonal differences; $\phi(B)$ is a stationary autoregressive (AR) polynomial in B; $\theta(B)$ is an invertible moving average (MA) polynomial in $B$. For seasonal series, the polynomials typically have a "multiplicative" structure. If $n$ denotes the number of observations per year in TRAMO, the polynomials in B can be factorized as

$$\delta(B) = (1-B)^d(1-B^s)^{d_s} = \nabla^d \nabla_s^{d_s}$$

where $\nabla$ and $\nabla_s$ are the regular and seasonal differences; and

$$\phi(B) = \phi_p(B)\,\Phi_{p_s}(B^s) = \left(1 + \phi_1 B + \dots + \phi_p B^p\right)(1 + \Phi_1 B^s) \tag{4}$$

$$\theta(B) = \theta_q(B)\,\Theta_{q_s}(B^s) = \left(1 + \theta_1 B + \dots + \theta_q B^q\right)(1 + \Theta_1 B^s) \tag{5}$$

Stationarity and invertibility imply that all of the roots of the polynomials in B on the right-hand side of (4) and (5) lie outside the unit circle. In what follows, variable $x_t$ will be assumed to be centred around its mean and the general expression for the model will be the ARIMA $(p, d, q)(p_s, d, q_s)_s$ model:

$$\phi_p(B)\,\Phi_{p_s}(B^s)\nabla^d\nabla_s^{d_s}x_t = \theta_q(B)\,\Theta_{q_s}(B^s)a_t \tag{6}$$

The regression variables can be defined by the user or generated automatically through the seasonal adjustment software. The generated variables are calendar variables, such as trading day (with two possible specifications), Easter and the leap year effect, intervention variables and outliers. Intervention variables reflect the dynamic patterns generated by dummy variables (and sequences thereof) aimed at capturing some known special effect. Outliers are atypical observations that distort the normality assumption for residuals, and need not be associated with known effects. As explained earlier, three types of outliers are considered (by default): additive outliers; transitory changes; and level shifts.

By default, TRAMO performs pre-tests for the log/level transformation as well as for the possible presence of calendar effects. The process is followed by an automatic model identification along with automatic outlier detection and correction, an estimation of the reg-ARIMA model using the exact maximum likelihood method, missing values interpolation and the and forecasting of time series data.

## 3.2 Automatic model identification in the presence of outliers

The algorithm iterates between the following two stages to identify the best model.

1. **Automatic outlier detection and correction**. This procedure is based on Tsay (1986) and Chen and Liu (1993) with some modifications. First, it assumes that a proper ARIMA model has been correctly identified for the series and uses this model to detect and correct the series for outlier effects. This method uses some form of ‘one by one’ as well as a joint outlier detection.

2. **Automatic model identification**. This consists of two steps: first, the differencing orders for the ARIMA model (polynomial $\delta(B)$) are automatically obtained. Second, it preforms the automatic identification of an ARMA model, i.e. $\phi_p(B)$, $\Phi_{p_s}(B^s), \theta_q(B)$ and $\Theta_{q_s}(B^s)$, for the differenced series, and is corrected for all outliers and other regression effects, if any.

For seasonal series, the default model in Stage 1 is the so-called “airline model” (Box-Jenkins, 1970), provided by the equation:

$$\nabla\nabla_s x_t = (1+\theta B)(1+\theta B^s)a_t, \tag{7}$$

i.e. the IMA (0,1,1)(0,1,1)$_s$ model, which is very flexible. It entails many other models and is very often applied in practice. For non-seasonal series, the default model is provided by the equation

$$\nabla x_t = (1 + \theta B)a_t + \mu, \tag{8}$$

i.e. the IMA (1,1) plus the mean model.

Identification of the ARIMA model is performed based on the time series corrected for outliers and regression effects. To modify the model, the automatic detection and outlier correction will begin in Stage 1.

### 3.3 System architecture

**Figure 4: Cycle of seasonal adjustment process in UNIDO Statistics**

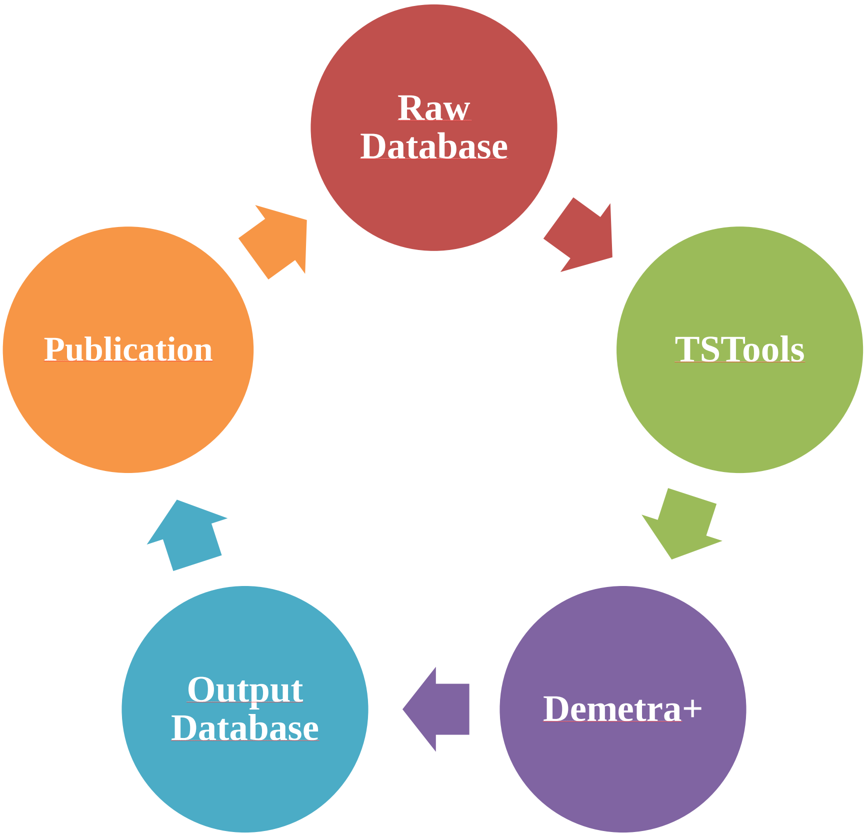
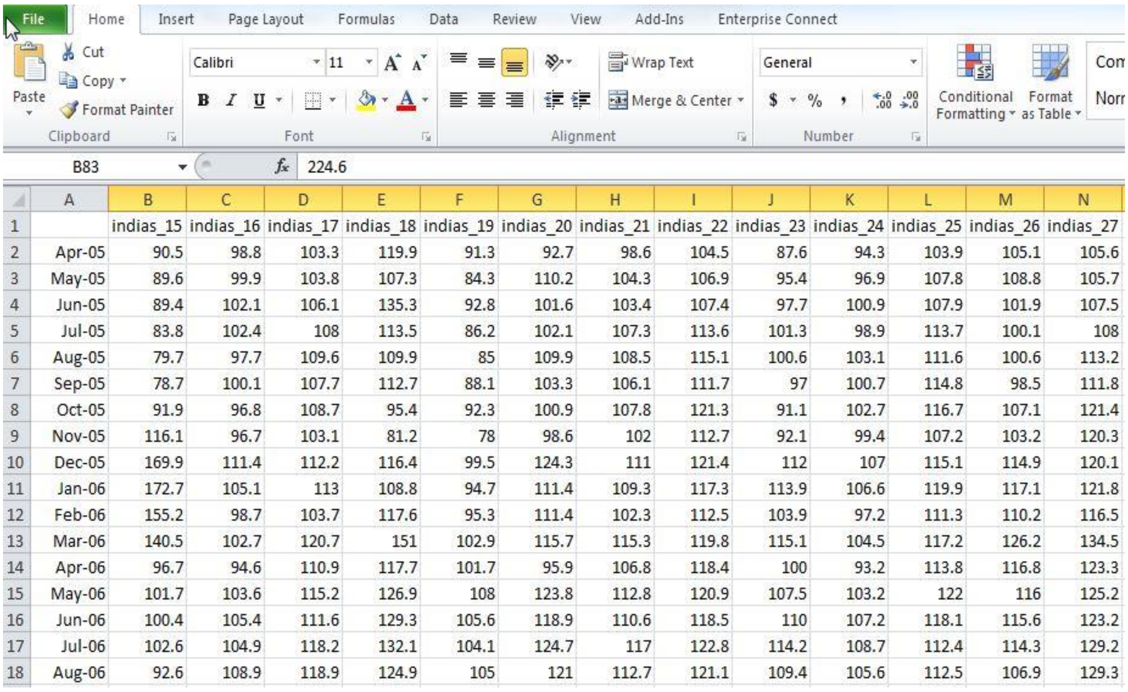


Figure 4 is a simple illustration of the seasonal adjustment of the sub-annual IIP indices procedure performed by UNIDO Statistics. The cycle starts by establishing and/or constructing an appropriate database of row, not seasonally adjusted, time series (data preparation step). Data should be collected using a special format such as Excel data sheets. Once the database has been prepared, the file will be considered a secure field which should not be modified, renamed or placed in another directory. This file contains the data in a proper format and can be imported to

Demetra+ as well as to TSTools. The first cell of the Excel sheet (namely A1) should remain empty [see Figure 5]. Neither a header note, nor an explanation should be added in this first cell. Then, missing the first cell, the first column includes the data identifiers (in case the user prefers a vertical representation of data), i.e. dates corresponding to the IIP data. The data identifier cells should be formatted to true calendar dates using Excel options. The nature of the data within the cells is important, hence, an inappropriate format in one cell will lead to an importing error in Demetra+. The next columns contain the IIP data relevant to the identifiers.

**Figure 5: Sample illustration of database with more than one time series**

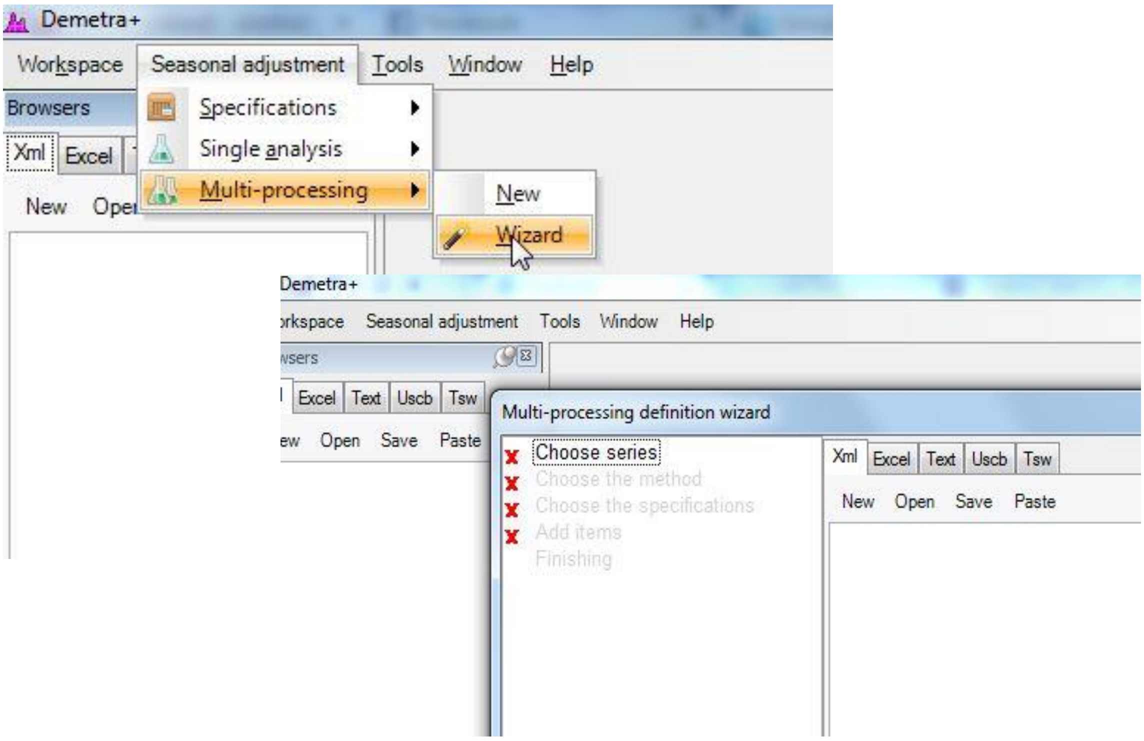


B83 224.6

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | indias_15 | indias_16 | indias_17 | indias_18 | indias_19 | indias_20 | indias_21 | indias_22 | indias_23 | indias_24 | indias_25 | indias_26 | indias_27 |
| 2 | Apr-05 | 90.5 | 98.8 | 103.3 | 119.9 | 91.3 | 92.7 | 98.6 | 104.5 | 87.6 | 94.3 | 103.9 | 105.1 | 105.6 |
| 3 | May-05 | 89.6 | 99.9 | 103.8 | 107.3 | 84.3 | 110.2 | 104.3 | 106.9 | 95.4 | 96.9 | 107.8 | 108.8 | 105.7 |
| 4 | Jun-05 | 89.4 | 102.1 | 106.1 | 135.3 | 92.8 | 101.6 | 103.4 | 107.4 | 97.7 | 100.9 | 107.9 | 101.9 | 107.5 |
| 5 | Jul-05 | 83.8 | 102.4 | 108 | 113.5 | 86.2 | 102.1 | 107.3 | 113.6 | 101.3 | 98.9 | 113.7 | 100.1 | 108 |
| 6 | Aug-05 | 79.7 | 97.7 | 109.6 | 109.9 | 85 | 109.9 | 108.5 | 115.1 | 100.6 | 103.1 | 111.6 | 100.6 | 113.2 |
| 7 | Sep-05 | 78.7 | 100.1 | 107.7 | 112.7 | 88.1 | 103.3 | 106.1 | 111.7 | 97 | 100.7 | 114.8 | 98.5 | 111.8 |
| 8 | Oct-05 | 91.9 | 96.8 | 108.7 | 95.4 | 92.3 | 100.9 | 107.8 | 121.3 | 91.1 | 102.7 | 116.7 | 107.1 | 121.4 |
| 9 | Nov-05 | 116.1 | 96.7 | 103.1 | 81.2 | 78 | 98.6 | 102 | 112.7 | 92.1 | 99.4 | 107.2 | 103.2 | 120.3 |
| 10 | Dec-05 | 169.9 | 111.4 | 112.2 | 116.4 | 99.5 | 124.3 | 111 | 121.4 | 112 | 107 | 115.1 | 114.9 | 120.1 |
| 11 | Jan-06 | 172.7 | 105.1 | 113 | 108.8 | 94.7 | 111.4 | 109.3 | 117.3 | 113.9 | 106.6 | 119.9 | 117.1 | 121.8 |
| 12 | Feb-06 | 155.2 | 98.7 | 103.7 | 117.6 | 95.3 | 111.4 | 102.3 | 112.5 | 103.9 | 97.2 | 111.3 | 110.2 | 116.5 |
| 13 | Mar-06 | 140.5 | 102.7 | 120.7 | 151 | 102.9 | 115.7 | 115.3 | 119.8 | 115.1 | 104.5 | 117.2 | 126.2 | 134.5 |
| 14 | Apr-06 | 96.7 | 94.6 | 110.9 | 117.7 | 101.7 | 95.9 | 106.8 | 118.4 | 100 | 93.2 | 113.8 | 116.8 | 123.3 |
| 15 | May-06 | 101.7 | 103.6 | 115.2 | 126.9 | 108 | 123.8 | 112.8 | 120.9 | 107.5 | 103.2 | 122 | 116 | 125.2 |
| 16 | Jun-06 | 100.4 | 105.4 | 111.6 | 129.3 | 105.6 | 118.9 | 110.6 | 118.5 | 110 | 107.2 | 118.1 | 115.6 | 123.2 |
| 17 | Jul-06 | 102.6 | 104.9 | 118.2 | 132.1 | 104.1 | 124.7 | 117 | 122.8 | 114.2 | 108.7 | 112.4 | 114.3 | 129.2 |
| 18 | Aug-06 | 92.6 | 108.9 | 118.9 | 124.9 | 105 | 121 | 112.7 | 121.1 | 109.4 | 105.6 | 112.5 | 106.9 | 129.3 |

After completing the preparation of available row data, an interim effort including data compilation using an additional tool, TSTools, is carried out. UNIDO Statistics is often faced with the problem of short time series. In such cases, there are insufficient observations to perform seasonal adjustments and forecasts. Seasonal adjustment methods require a minimum of 36 observations for monthly and 16 observations for quarterly time series. In most cases, annual data for preceding years are available while quarterly/monthly data are missing. When the length of monthly/quarterly time series is insufficient to perform seasonal adjustment, TSTools is a useful means to expand the time series. TSTools software performs so-called temporal disaggregation of annual data using available sub-annual information, and disaggregates annual data into monthly/quarterly format. Using this method may increase the number of short-term observations and facilitate the performance of a seasonal adjustment.

**Figure 6: Creating a multi-processing specification to adjust more than one time series**



Following the data preparation steps (using the raw database and TSTools), the seasonal adjustment in Demetra+ is performed to obtain seasonal adjusted indices. In this step, a workspace is created by the user. This workspace is saved and used for the entire year (based on UNIDO Statistics’ revision policy). The specifications are defined and saved in the workspace in the first quarter of each year and is used for the next quarters until the end of each year.

While creating the specification [see Figure 6] after having selected the series, the system asks the user to choose between two options: TRAMO/SEATS and X12. As discussed earlier, here we use TRAMO/SEATS. We then have to select the specification. The details of each available option are described in Table 2.

The RSA5 specification is the most comprehensive option in automatic model identification. By choosing this option, we let the algorithm, as defined in the software, select the best model. This accelerates the process by reducing manual intervention by the user. The results using RSA5 are often very reliable. Model modification by user is, however, still possible.

The fourth step involves exporting the output (seasonally adjusted time series and components) in a proper format for further compilations. Demetra+ provides the user with a variety of data export formats. As regards the nature of the usage, three types of data can be published after seasonally adjusting the time series: original data, calendar adjusted data, and seasonal and calendar adjusted data.

**Table 2: Available specification options in TRAMO/SEATS**

| **Specification** | **Settings** |
|---|---|
| RSA0 | Use of default Airline model, producing models in levels |
| RSA1 | As RSA3, but the default Airline model is always used |
| RSA2 | As RSA4, but the default Airline model is always used |
| RSA3 | The software tests the log/level specification, interpolates missing observations (if any), performs automatic model identification and outlier detection. Three types of outliers are considered: additive outliers, transitory changes and level shifts. The level of significance is set by the programme and depends on the length of the series. The full model is estimated by the exact maximum likelihood, and forecasts of the series up to a one-year horizon are computed. The model is decomposed and forecasts are obtained for each component (trend, seasonal, irregular and (possibly) transitory). If the model does not accept an admissible decomposition, it is replaced by a decomposable one. |
| RSA4 | As RSA3, but a pre-test is carried out for the presence of a trading day, leap year and Easter effects, with one parameter specification for trading days (working/ non-working days). |
| RSA5 | As RSA4, but the trading day specification uses 6 parameters (effect of all weekdays are pre-tested). |

The cycle illustrated in Figure 2 is continued by updating the raw data in the next quarter. In our experience, modifications are made manually. The major modifications are summarized in Table 3.

**Table 3: Modification made manually in the process of seasonal adjustment of IIP**

| **Modification made** | **Details** |
|---|---|
| Shortening the time span (modifications can be made using the "Transformation" option) | To set annual totals, autocorrelation and normality problems as well as "zero" observations. |
| Imposing the default (Airline) model (0,1,1)(0,1,1)<br>(set the "Auto modelling" option to "Disabled") | In rare cases, the ARIMA model defined by TRAMO shows residual seasonality. This occurs when the identified model cannot capture seasonality. The Airline model is a good benchmark model in such cases.<br>The automatically selected ARIMA model may produce irregular components that are correlated with seasonal components. The Airline model's flexibility may offer a solution in this case as well. |
| Imposing trading day effect<br>(set the "Pre-test" option under "Trading days" to "False") | Although the software pre-tests the presence of the trading day effect, it sometimes does not capture almost significant effects. The residual trading day graph helps experts decide whether trading day effects should be imposed manually or not. |
| No trading day effect<br>(set the "Type" option under "Trading days" to "None") | If the software's main results indicate "definition error", the trading day effect should be removed from the model. The same holds for "seas-irr correlation" error. |
| Intervention effect (seasonal outlier) | When the time series shows seasonality at certain time intervals but not for the entire time series, the automatically identified model failed to capture seasonality. The model should be identified by the expert using intervention variables. |
| Decreasing the critical value to detect outliers<br>(set "Default critical value" to "False" and change the critical value) | Demetra+ defines a default critical value to detect outliers. The expert can change the value manually if significant outliers are evident in the graphs. |

Demetra+ automatically offers the option of refreshing the workspace when new data become available. Refreshment strategies are explained in the next section.

## 3.4 Revision policy

Revisions of seasonally adjusted data are done for two main reasons. First, seasonally adjusted data are modified due to a revision of the unadjusted (raw) data. Revisions of unadjusted data may be attributable to an improved information set (in terms of coverage and/or reliability).

Second, revisions of seasonally adjusted data can also be made due to a better estimate of the seasonal pattern based on new information provided by new unadjusted data and due to the characteristics of the filters and procedures removing seasonal and calendar components. Revisions are generally accepted if they are based on new information. However, in seasonal adjustment the case may be that just one observation more results in revisions of the seasonally adjusted data for several years, which can confuse the user.

The challenge is to find a balance between the need for the best possible seasonally adjusted data, especially towards the end of the time series, and to avoid unnecessary revisions that may later be reversed (the trade-off between the precision of seasonally adjusted data and their stability over time).

Prior to developing a revision policy, consideration needs to be given to the needs of users and available resources to implement the policy. The policy should at the very least address the following: the frequency and relative size of revisions due to seasonal adjustment; the precision of the seasonally adjusted data, the time period over which the raw data have been revised and the relationship between the timing of publication of revisions to the seasonally adjusted data and publication of the revisions to the raw data.

The revision policy must be as coherent and transparent as possible and should not lead to the publication of sub-optimal seasonally adjusted data, which could mislead users who conduct economic assessments.

There are several options for revisions:

- Revise seasonally adjusted data in accordance with a well-defined and publically available revision policy and release calendar;
- Revise both raw and seasonally adjusted data between two consecutive official releases of the release calendar;
- Revise seasonally adjusted data only once a year independently of any revisions of past raw data;
- Revise seasonally adjusted data once a year if past raw data do not change when a new observation is added, or revise seasonally adjusted data whenever past raw data are revised;
- Do not use any official release calendar and/or perform revisions on an irregular basis and/or do not revise at all.

The best applicable method is defined here. Revisions to seasonally adjusted data should be published in accordance with a coherent, transparent and officially published revision policy and release calendar, which is aligned with the revision policy and the revision calendar for the unadjusted data. Seasonally adjusted data should not be revised more often than releases of raw /unadjusted data. The public should be informed about the average revisions of important seasonally adjusted macroeconomic variables, which have been observed in the past.

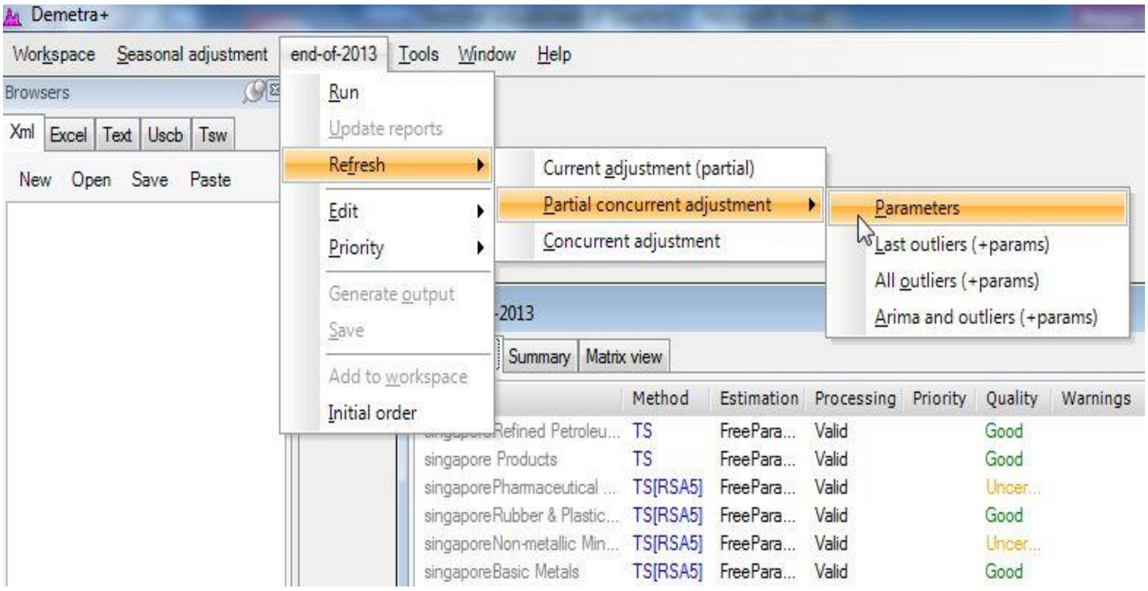
Demetra+ offers three types of revision possibilities:

- Current adjustment → adjustment with fixed specification while user defined regression variables can be updated
- Semi-concurrent revision → re-estimates respective parameters and factors each time new or revised observations become available
- Concurrent adjustment → adjustment performed without any fixed specifications.

When past data are revised for less than two years and/or new observations become available, a partial concurrent adjustment (semi-concurrent revision) is preferred in order to take the new information into account and to minimize the extent of the revision(s) because of the seasonal adjustment process. However, if the seasonal component is sufficiently stable, a controlled current adjustment could be considered to minimize the frequency of revisions. In that case, a full review of all seasonal adjustment parameters should be carried out at least once a year. When revisions covering two or more years occur, filters, outliers and regression parameters have to be re-identified and re-estimated. The revision period for seasonally adjusted data must at least cover the raw data revision period. Due to the property of filters, it is usually acceptable to revise the seasonally adjusted data from a point of 3-4 years before the beginning of the revision period of the unadjusted data; earlier data should be frozen.

In UNIDO Statistics, the new specification is defined when all the sub-annual IIP data become available. The models are revised based on a semi-concurrent revision [see Figure 7] for each quarter when new data become available, until the fourth quarter of the given calendar year.

**Figure 7: How to revise the model based on partial concurrent adjustment**



| | Method | Estimation | Processing | Priority | Quality | Warnings |
|---|---|---|---|---|---|---|
| Refined Petroleu... | TS | FreePara... | Valid | | Good | |
| singapore Products | TS | FreePara... | Valid | | Good | |
| singaporePharmaceutical ... | TS[RSA5] | FreePara... | Valid | | Uncer... | |
| singaporeRubber & Plastic... | TS[RSA5] | FreePara... | Valid | | Good | |
| singaporeNon-metallic Min... | TS[RSA5] | FreePara... | Valid | | Uncer... | |
| singaporeBasic Metals | TS[RSA5] | FreePara... | Valid | | Good | |

## 3.5 Costs and risks

Seasonal adjustment is a time consuming procedure, and significant computer/human resources must be earmarked for this task. Unsuitable or low-quality seasonal adjustment can generate misleading results and increase the probability of false signals (credibility effects). The presence of residual seasonality as well as over-smoothing, are concrete risks that can negatively affect the interpretation of seasonally adjusted data. Series that do not show the presence of seasonality should not be seasonally adjusted.

# 4. Application

Since the first quarter of 2013, nearly 335 time series of IIP at the sector level have been seasonally adjusted by UNIDO Statistics. To improve the quality of seasonally adjusted data for time series that are shorter than three years, temporal disaggregation is performed. According to preliminary results, 85 percent of the time series had significant underlying seasonality and/or trading day effects. The seasonally adjusted and original IIP data for two sample countries are illustrated in Figures 8 and 9.

As can be seen in the figures, the seasonally adjusted results do not show “normal” and recurring events in the time series; they provide an estimate for what is new in the series which is the ultimate goal of seasonal adjustment. It is easy to identify the seasonal pattern of the original time series (see Figure 9). The production index drops repeatedly in the 3rd quarter (summer) in almost every year. This could be attributable to changes in working hours during the summer or to the country's weather conditions, which affects productivity. Fluctuations due to exceptionally strong or weak seasonal influences will continue to be visible in the seasonally adjusted series (e.g. in Figure 9, the outlier that occurred in QI 2008 can be seen in the SA time series as well). Other random disruptions and unusual movements that are readily understandable in economic terms (for example, the consequences of economic policy, large-scale orders or strikes) will also continue to be visible.

**Figure 8: Seasonally adjusted and original series IIP**

**Figure 9: Seasonally adjusted and original series IIP**

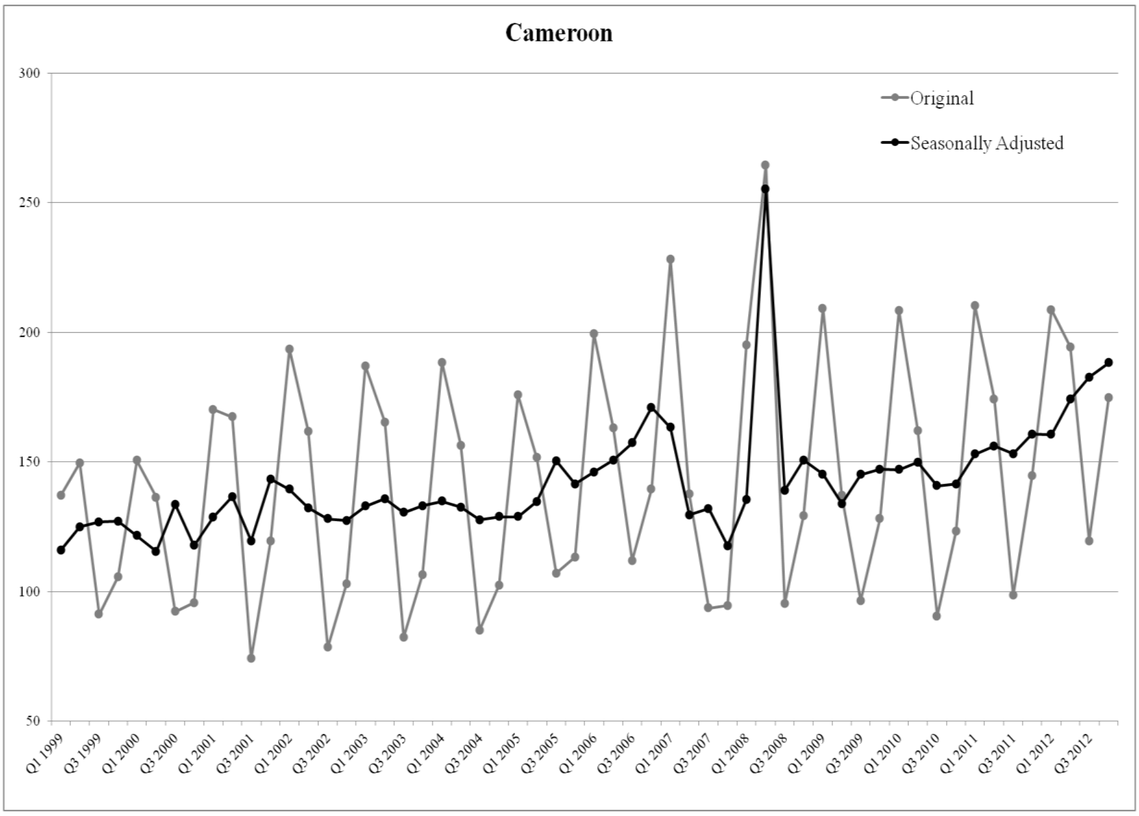


Table 4 shows the difference in growth rate figures before and after performing seasonal adjustment. This provides a better picture of why seasonal adjustment is necessary and how seasonal adjustment helps reveal the real movements and turning points in manufacturing production, which may be impossible or difficult to identify due to seasonal movements. For instance, the original 2012 data for Cameroon indicates a decline of 6.9 percent in manufacturing production when removing the seasonal effect; the SA data, on the other hand, shows a significant growth of 8.4 percent.

## 5. Quality of the original data

It is important to look at the data and graph of the original time series the reviewer seeks to adjust before running Demetra+ for seasonal adjustment. Graphic illustration of the time series is a powerful tool to identify quality issues. As mentioned earlier, Demetra+ requires at least three years of data to compute a seasonal adjustment and often more than four years of data to obtain an adequate adjustment and a more reliable forecast.

**Table 4: IIP percentage change**

| | QII 2012 to QI 2012 | | QII 2011 to QI 2011 | |
|---|---|---|---|---|
| | Russia | Cameroon | Russia | Cameroon |
| Original | -1.39% | -6.86% | -15.48% | -17.17% |
| SA | -0.66% | 8.43% | -8.51% | -0.66% |

Series with too many zero values in a row will cause estimation problems for Demetra+ and the software may estimate negative values for those units or may simply send an error message that the series cannot be seasonally adjusted. Large outliers may also cause problems for the adjustment if they are not properly accounted for. The user should look for changes in the time series pattern and at their extent to determine whether the span of data to be used for seasonal adjustment should be modified.

Short time series should also be treated differently. A temporal of the annual data disaggregation using TSTools could be a solution. However, as this does not precisely match the real pattern of the time series, it might affect the quality of seasonally adjusted data. Before performing the adjustment process, time series with fewer than three years of data need to be identified. If the annual data are not available either, the process should be stopped; the series is not long enough for seasonal adjustment.

Series with possible outlier values should be identified as well (additive outliers, level shifts and temporary changes). The user should verify that the outliers are valid and do not indicate a problem.

It may occasionally be better to shorten series with obvious changes in the pattern. A more consistent pattern in the time series often improves the seasonal adjustment. It is not always easy to identify changes, but graphs of the original series and year-by-year growth rate graphs are useful for this purpose.

## 6. Discussion

UNIDO Statistics began performing seasonal adjustments on the secondary index of industrial production at the two-digit level at international level. However, it is highly recommended to perform a seasonal adjustment exercise at country level. This will provide more precise results and prevent many unknown constraints which may emerge at the international level.

Users can choose between one of the seasonal adjustment approaches for each time series under review, but once a method has been selected, it is not advisable to switch. This would lead to pointless revisions and only confuses data users. The same holds for the revision policy. Before initiating a SA, the revision policy should be clear and the revision strategies made available to the data users.

Moreover, the publication policy should be clear in a way that when seasonality is present and can be identified, time series should be made available in a seasonally adjusted format. It is recommended for the method and software used to be explicitly mentioned in the metadata accompanying the series.

Countries with no seasonal adjustment experience are encouraged to compile, maintain and update their national calendars or, as a minimum alternative, to supply a list of past public holidays including, where possible, information on compensation holidays. Moreover, the calendar for the year t+1 or the corresponding holiday list should be provided.

One common misconception is that seasonal adjustment hides outliers present. This is not the case: if some unusual event occurs, this information must be analysed, and outliers are included in the seasonally adjusted time series. However, users of seasonally adjusted data should be aware that the data’s usefulness for econometric modelling purposes needs to be carefully considered.

The user has to decide whether to seasonally adjust an aggregate such as aggregated IIP at total and sector level directly or to sum up its seasonally adjusted components. If the main purpose is to preserve the relationship between data, the indirect approach is more appropriate. For time series that have very similar seasonal components, direct adjustment would suffice, noting that summing up the series together will first reinforce the seasonal pattern while allowing the cancellation of some noise in the time series.

## References

Box, G.E.P., and Jenkins, G.M., 1976, Time Series Analysis, Forecasting and Control, Holden-Day, San Francisco.

Chen, C. and Liu, L., 1993, "Joint Estimation of Model Parameters and Outlier Effects in Time Series", Journal of the American Statistical Association, 88, 284-297.

"ESS Guidelines on Seasonal Adjustment" 2009 Edition, Eurostat, European Commission, ISBN 978-92-79-12307-8, active link on April 24, 2013 http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-RA-09-006/EN/KS-RA-09-006-EN.PDF

"Practical Guide to Seasonal Adjustment with Demetra+", UNECE, 2012, active link on April 24, 2013 http://www.unece.org/fileadmin/DAM/stats/publications/Practical_Guide_to_Seasonal_Adjustment_final_web.pdf

"Reliability of the Tramo-Seats Automatic Identification of Regression-Arima Models" Unpublished manuscript, 31.08.2012. Active link May 12, 2013 http://www.cros-portal.eu/sites/default/files//JDemetra%2B%20user%20manuals_0.zip

"Seasonal Adjustment", European Central Bank, 2012. Active link on May 12, 2013 http://www.ecb.int/pub/pdf/other/statseasonaladjustmenten.pdf

"Time Series Analysis – Basics" Australian Bureau Statistics. Active link on May 12, 2013 http://www.abs.gov.au/websitedbs/D3310114.nsf/home/Time+Series+Analysis:+The+Basics

Tsay, R., 1986, "Time Series Model Specification in the Presence of Outliers", Journal of the American Statistical Association, 81, 132-141.

V.15-01458—March 2015—275

UNITED NATIONS
INDUSTRIAL DEVELOPMENT ORGANIZATION

Vienna International Centre · P.O. Box 300 · 1400 Vienna · Austria
Tel.: (+43-1) 26026-0 · E-mail: unido@unido.org
www.unido.org