UNITED NATIONS
INDUSTRIAL DEVELOPMENT ORGANIZATION

Seasonal Adjustment
of Indices of Industrial Production

Methodological Note

October 2020

Contents

1 Introduction **2**

2 Theory of seasonal adjustment **3**

2.1 Time series and its components . . . 3

2.1.1 Seasonal and calendar effects . . . 3

2.1.2 Trend, irregular and other effects . . . 5

2.1.3 Composition models . . . 6

2.1.4 Outliers . . . 6

2.2 Seasonal adjustment approaches . . . 8

2.2.1 Types of seasonal adjustment approaches . . . 9

2.2.2 Reg-ARIMA models in TRAMO/SEATS . . . 9

3 Calculation of seasonal adjustment **13**

3.1 Adjusting with JDemetra+ Software . . . 13

3.2 Data preparation and requirements . . . 14

3.3 Data handling . . . 16

3.4 Examples . . . 19

3.5 Revision policy . . . 20

3.6 Data quality and costs . . . 22

3.7 Data presentation . . . 23

4 Conclusion **25**

References **26**

1 Introduction

The sub-annual Index of Industrial Production (IIP) is one of the most important short-term indicators measuring the development of industrial output volume. It is an important indicator for detecting business cycle's turning points as early as possible. This allows planners, decision makers and the business community at large to detect possible changes in the economy in order to take appropriate and timely policy measures.

However, these statistics are often influenced by seasonal as well as calendar effects (e.g. moving holidays or deviating fiscal years), which can mask relevant short- and long-term movements of the time series and prevent a clear understanding of economic phenomena. In this way, the seasonally adjusted results do not show "regular" and repeating events, thus revealing the "news" contained in the time series ([10, p. 102]).

In summary, the main reasons for seasonal adjustment of IIPs for UNIDO Statistics are:

- providing more reliable short-term forecasts of industrial activities;
- comparing industrial production growth rates from different countries;
- comparing growth rates with the previous quarter;
- revealing real movements and turning points in manufacturing production.

UNIDO Statistics publishes seasonally adjusted quarterly data[1] and a corresponding quarterly report on World Manufacturing Production[2] since 2011. Furthermore, UNIDO Statistics started publishing seasonally adjusted monthly data[3] in the beginning of 2020. Additional methodological information regarding index compilation can be accessed through the respective methodological guidelines [9].

[1] https://stat.unido.org/database/Quarterly%20IIP

[2] http://www.unido.org/resources/statistics/quarterly-report-on-manufacturing.html

[3] https://stat.unido.org/database/Monthly%20IIP

2 Theory of seasonal adjustment

The following section discusses important concepts of seasonal adjustment, such as seasonal and calendar effects, outliers as well as different seasonal adjustment approaches.

2.1 Time series and its components

A time series is a set of statistics, collected at regular intervals, e.g. quarterly or monthly IIP or monthly data on unemployment. Usually, the intent is to describe and summarize the data, apply a model and detect a (possible) discernible pattern with the intention of short-term modelling as a base for policy making. Time series data diverges from other data due to its time dependence and correlation, so that many statistical methods are not applicable because of the common assumption of independence.

An economic time series records the direction and turning points of economic activities. In such cases, data undergo regular and fairly predictable changes, recurring every calendar year (so-called periodic fluctuations, such as Christmas). These regular patterns are called seasonality.

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

- Seasonal and calendar effects (S): cyclical fluctuations related to the calendar (including moving holidays and working day effects).
- Trend (T): long-term movements at the level of the series.
- Cycles (C): cyclical fluctuations longer than a year (such as business cycles).
- Irregular (I): (short-term) unpredictable or other random fluctuations (such as strikes or unusual weather situations).

2.1.1 Seasonal and calendar effects

Seasonal and calendar effects in economic 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. The example in figure 1 shows the original as well as seasonally adjusted IIP of Colombia together with the trend series. The calendar-related and regular movements of the original data are excluded in the seasonally adjusted as well as the trend data.

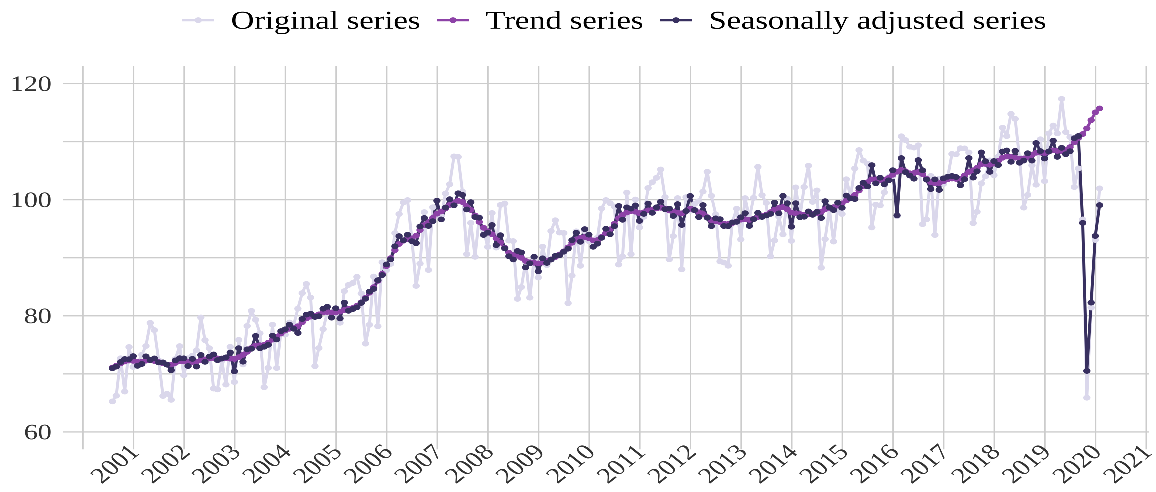

Figure 1: Comparison of seasonally adjusted and unadjusted IIP of Colombia

The possible causes of seasonal effects can be classified into three groups:

- natural factors,
- administrative or legal rhythms or
- social/cultural/religious traditions (e.g. fixed holidays or timing of vacations).

Other calendar-related effects are usually included in the seasonal effects, although they are sometimes considered as separate components of time series. Moving holidays and the varying number of holidays of different countries (such as Easter holidays, Ramadan or the Chinese New Year) need to be included in the model. Another calendar-related effect refers to the leap year, where another (trading or weekend) day is added every four years. Working day adjustment assumes a difference between working and non-working days (Saturdays and Sundays). Most financial time series are influenced by this effect. On the other hand, the trading days effect influences many real industrial activities and refers not also to the difference between working and non-working days, but also to the varying economic activity between the different working days. In the example of table 1, the counts of all weekdays (except Wednesdays) vary between October and November 2020, which could have a significant influence on economic time series. Furthermore, the trading days adjustment assumes no economic activity on Sundays, which is usually modeled with six regression variables for the remaining weekdays.

	M	T	W	T	F	S	S
				1	2	3	4
	5	6	7	8	9	10	11
October 2020	12	13	14	15	16	17	18
	10	20	21	22	23	24	25
	26	27	28	29	30	31	
							1
	2	3	4	5	6	7	8
	9	10	11	12	13	14	15
November 2020	16	17	18	19	20	21	22
	23	23	24	25	26	27	28
	29	30					

Table 1: Trading days effect

2.1.2 Trend, irregular and other effects

The trend component (including cycles, also known as the trend-cycle component) 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.

The irregular or residual component is the time series' unpredictable component, which will remain after excluding the seasonal (including calendar-related effects) and the trend components from the original series. It captures the remaining unsystematic or unpredictable short-term fluctuations in terms of timing, impact as well as duration and includes errors of measurement and unusual events. The major causes of irregular effects include unseasonable (unexpected) natural disasters, strikes as well as sampling and non-sampling errors.

2.1.3 Composition models

In case of index of industrial production, we find:

$$Z_t = \text{IIP at time t}$$

The idea behind seasonal adjustment is to create separate models for all the time series' components and combine them additively:

$$Z_t = S_t + T_t + I_t$$

where the seasonal adjusted series is derived as:

$$SA_t = Z_t - S_t \text{ or } SA_t = T_t + I_t$$

or multiplicatively:

$$Z_t = S_t \times T_t \times I_t$$

where

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

Seasonally adjusted time series (SA) only contain irregular (I) and trend-cycle (T) components. The additive decomposition model assumes independent behavior of the components of time series. In other words, the amount of seasonal and irregular variations does not change if 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. The degree of seasonal variation increases as the level of the components' trend rises. In the multiplicative model, the original time series is expressed as the product of trend-cycle (T), seasonal (S) and irregular (I) components.

2.1.4 Outliers

Outliers are extreme observations that deviate from the trend and fall outside the expected range of typical time series patterns. These abnormal values may occur, for instance, as a result of new policies, new types of taxes, extreme natural events or a closure of a significant manufacturer. These outliers contain valuable information about unusual events. They are an important part of the data and will remain visible in seasonally adjusted time series. However, outliers need to be identified to avoid distorting the estimation of the seasonal component. Large economic uncertainties (e.g. the economic crisis caused by COVID-19) have to be considered as outliers in the seasonal adjustment process.

The three most frequent outliers (see figure 2) are additive outliers, temporary changes[4] and level shifts.

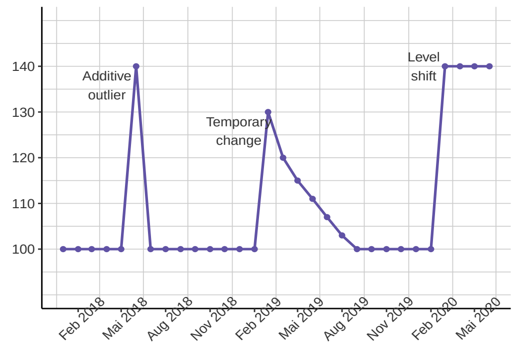

Figure 2: Frequent forms of outliers

An additive outlier (AO) affects a single observation. After this disruption, the series returns to its normal path as if nothing happened. The AO may be caused by a random effect, such as a strike or a short-term shock in the system. For instance, a pre-announced price rise could drastically increase sales before the price change is even introduced.

Temporary changes (TC) are spikes that take several periods to disappear gradually. An example are deviations from average monthly weather conditions. If the weather changes drastically, energy consumption may rise or fall.

A level shift (LS) refers to a more permanent change in the time series' level often as a result of changes in economic behavior, social traditions or legislation. Level shifts change the level of the time series, but do not modify seasonal behavior. It may also occur due to changes in concepts and definitions of the survey population or compilation methods, which should be avoided to preserve the comparability of the series.

[4] Also referenced as transitory changes.

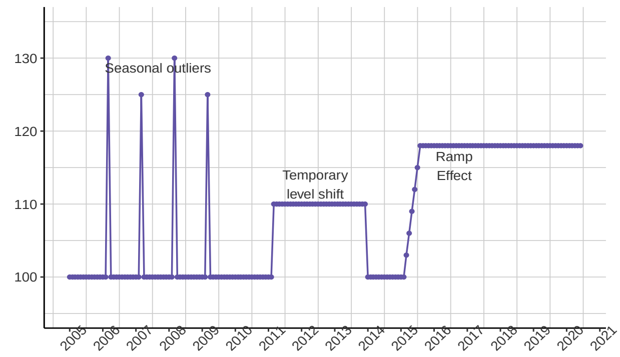

Figure 3: Other forms of outliers

Additionally, there are other types of outliers (see figure 3), such as seasonal outliers (SO), temporary level shifts or ramp effects. In contrast to the above mentioned outlier types, they are not detected automatically within standard seasonal-adjustment procedures (see chapter section 2.2.2), but can be manually identified by the user. Seasonal outliers are determined by a sudden drop or rise occurring in the same period regularly, but not every year with a similar magnitude, which can for example be explained by different weather situations, such as cooler or hotter summer periods than the average. Temporary level shifts, on the other hand, relate to a shift in the data but, unlike level shifts, only for a limited time, after which the series returns to its preceding level. This might be caused by an economic crisis in the specific region. Ramp effects indicate a linear increase or decrease of the level of a series during a certain period, with the same triggers as level shifts.

While identifying outliers, it needs to be considered that different outlier types may impact different components of the time series. For instance, a LS is related to the trend cycle component, whereas an AO or a TC are assimilated to the irregular component. The modelling of extreme situations, such as the economic crisis of 2020, would most likely require level shifts or temporary changes as economists point to a longer global economic downturn followed by a slow recovery.

2.2 Seasonal adjustment approaches

Seasonal adjustment approaches can generally be classified into two main groups: model-based and filter-based methods, which will be further explained below.

2.2.1 Types of seasonal adjustment approaches

The model-based approach specifies explicit statistical models of the trend, seasonal and irregular components of the classical decomposition and derives the different observations in accordance to this models. Examples of model-based methods include TRAMO/SEATS and STAMP. As TRAMO/SEATS is the chosen method for seasonal adjustment within UNIDO Statistics, it is further described below.

The filter-based approach, on the other hand, is based on the 'ratio to moving average' procedure and applies 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. The main procedure usually 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 or X12-ARIMA (using regARIMA models for forecasts, backcasts and pre-adjustments).

2.2.2 Reg-ARIMA models in TRAMO/SEATS

TRAMO (Time Series Regression with ARIMA[5] Noise, Missing Observations and Outliers) and SEATS (Signal Extraction in ARIMA Time Series) were developed by Víctor Gómez and Agustín Maravall [4] at the Bank of Spain. The two programs are used by different statistical agencies, including Eurostat and the European Central Bank. TRAMO and SEATS are parametric methods and 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. Further information can be found for example in [7] or [3].

[5] Auto-Regressive Integrated Moving Average

2.2.2.1 Theoretical background

Let z represent a time series observation

$$z = (z_{t_1}, z_{t_2}, ..., z_{t_m})$$

where $1 = t_1 < t_2 < ... < t_m = T$ are the time points. Missing observations are included and a log-transformation of the original observations might be used to provide the stationarity of the data. The Reg-ARIMA model can be derived as

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

where $y_t' = (y_{1_t}, ..., y_{n_t})$ is a matrix with n regression variables (related to outliers, calendar effects or other user-defined variables) and $\beta = (\beta_1, ..., \beta_n)'$ is the vector of corresponding regression coefficients. The variable x_t follows a (possibly non-stationary) ARIMA model and represents the stochastic component, whereas β represents the deterministic one. The ARIMA model for x_t can be written as

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

where B is the backshift operator, such that $B^j z_t = z_{t-j}$. Moreover, v_t is the stationary transformation of x_t and $\delta(B)$ is a non-stationary autoregressive (AR) polynomial containing regular and seasonal differences.

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

where μ_v is the mean of v_t and in practice it's an estimated regression parameter. $\phi(B)$ is a stationary AR polynomial in B and $\theta(B)$ is an invertible moving average (MA) polynomial in B. a_t refers to the white noise error (mean = 0, and constant variance) and can be interpreted as the forecasting error of x_t. For seasonal series, the polynomials typically follow a "multiplicative" structure, leading to polynomials in B as

$$\delta(B) = (1 - B)^d (1 - B^s)^{d_s} = \nabla^d \nabla_s^{d_s} \tag{4}$$

where s denotes the number of observations per year in TRAMO and ∇ and ∇_s are the regular and seasonal differences; and

$$\phi(B) = \phi_p(B) \Phi_{p_s}(B^s) = (1 + \phi_1 B + ... + \phi_p B^p)(1 + \Phi_1 B^s) \tag{5}$$

$$\theta(B) = \theta_q(B)\Theta_{q_s}(B^s) = (1 + \theta_1 B + ... + \theta_p B^q)(1 + \Theta_1 B^s) \tag{6}$$

where p is the number of regular AR terms ($<= 3$), p_s the number of seasonal AR terms ($<= 1$), q the number of regular MA terms ($<= 3$) and q_s the number of seasonal MA terms ($<= 1$). d is the regular differencing order and d_s the seasonal differencing order.

Stationarity and invertibility imply that all of the roots of the polynomials in B on the right-hand side of (5) and (6) lie outside the unit circle. In what follows, variable x_t will be assumed to be centered around its mean and the general expression for the model will be the ARIMA $(p, d, q)(p_s, d, q_s)_s$ model:

$$\phi_p(B)\Phi_{p_s}(B^s)\nabla^d\nabla_s^{d_s}x_t = \theta_q(B)\Theta_{q_s}(B^s)a_t \tag{7}$$

The regression variables (calendar variables as well as intervention variables and outliers) can be defined by the user or generated automatically through the seasonal adjustment software. Intervention variables reflect the dynamic patterns generated by dummy variables (and sequences thereof) aimed at capturing a-priori known special events (such as strikes or price devaluations). Outliers are simple forms of interventions variables and atypical observations that distort the normality assumption for residuals. Identifying the above-mentioned outlier types (see section 2.1.4) is crucial because Reg-ARIMA models are based on moving averages and therefore not robust and sensitive to extreme values. Additive outliers, temporary changes and level shifts are considered by default.

By default, TRAMO performs pre-tests for the log/level transformation of the series 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 (see below), an estimation of the regARIMA model using the exact maximum likelihood method, missing values interpolation and the forecasting of time series data.

2.2.2.2 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) [8] and Chen and Liu (1993) [2] 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” [1], provided by the equation:

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

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

$$\nabla x_t = (1 + \theta B)a_t + \mu, \tag{9}$$

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.

3 Calculation of seasonal adjustment

The following section describes requirements and the process of seasonal adjustment for IIP within UNIDO Statistics.

3.1 Adjusting with JDemetra+ Software

Since 2017, UNIDO Statistics achieves seasonal adjustment using the TRAMO/SEATS method in the JDemetra+ software with a partial concurrent revision policy (see section 3.5) when a new observation becomes available. A full review of all seasonal adjustment parameters is carried out once a year, usually at the beginning of the production year.

The JDemetra+[6] project/software 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) has been promoting the development of freely available JDemetra+ solutions for seasonal adjustment to be used within European statistical systems for several years. The development of the JDemetra+ 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 JDemetra+.

JDemetra+ is a family of modules on seasonal adjustment, which are based on the two leading algorithms in that domain, TRAMO/SEATS and X-13-ARIMA-SEATS (in short, X-13; more detailed information can be found in [11]). The literature does not propose any best method selection criteria between TRAMO/SEATS and X-13. TRAMO/SEATS (see chapter 2.2.2) is a parametric method and has a model-based facility to extract the underlying components. Therefore, the results of TRAMO/SEATS can be statistically explored. On the other hand, X-13 was developed by the U.S. Census Bureau and is a filter-based method, where in a first phase the regArima-model is estimated and afterwards the common X-11 method (using moving averages) together with SEATS is used for seasonal adjustment. Statistical evaluation is provided by the last phase, diagnostics, to measure the reliability of

[6] https://ec.europa.eu/eurostat/cros/content/softwareJDemetra_en

the results.

The comparison of the two methods is often difficult as their diagnostics focus on different aspects and their outputs take different forms. JDemetra+ provides a variety of graphical presentation and analysis tools for a better comparison of different methods. Therefore, a number of procedures for both methods has been re-written in JDemetra+, following an object-oriented approach. This could imply minor discrepancies in comparison to the original program in diagnostics or peripheral information, which should not, however, alter the general information provided by the algorithms. In any case, the original program's main results (i.e. the seasonally adjusted series) should not be affected by this solution.

UNIDO Statistics selected JDemetra+ and the TRAMO/SEATS method to produce seasonally adjusted IIPs mainly due to the following reasons:

- use of easy-to-interpret statistical tests;
- accurate automatic outlier detection and correction as well as automatic model identification;
- low probability of finding spurious seasonality; and,
- a simple and easy-to-use interface.

The version of JDemetra+ 2.2.2 (released on 15 January 2019) can be easily downloaded and is compatible with the latest Microsoft Office version. For working with Excel workbooks, the appropriate Microsoft Access DataBase Engine has to be installed. Detailed information on JDemetra+ can be found in the reference manual [6] or in the user guide [5].

3.2 Data preparation and requirements

The main data source of the sub-annual IIP published by UNIDO Statistics represents index numbers compiled and disseminated by National Statistical Offices (NSOs) through publications on their websites or databases.[7]

The IIP measures the growth of industrial production in real terms, free from price fluctuations. While annual industrial growth rates from national accounts generally refer to

[7] For the majority of European countries, IIP data are obtained from the Eurostat database, which contains data reported directly by NSOs

changes in manufacturing value added (MVA), i.e. output net of intermediate consumption, the sub-annual indices reflect the growth of gross output. Given the temporal nature of estimates, output growth provides the best approximation of value added growth, assuming that the input-output relationship is relatively stable during the observation period. To obtain internationally comparable IIP growth estimates from national data, certain methodological standards must be respected and implemented at the national level. These internationally accepted standards are introduced in the International Recommendations for the Index of Industrial Production (IRIIP) [10].

The index reference period is the period against which other periods are compared and for which the index is set equal to 100. The weight reference period is the period whose values serve as weights for the index. Starting with the reporting year 2020, the selected base period in UNIDO IIP publications is the year 2015. In case NSOs provide data with another base period, the data were rebased to fit the base year 2015 = 100.

Before starting the seasonal adjustment process, a decision needs to be made regarding the direct or indirect approach. In the direct approach, each time series is adjusted, even if this time series is an aggregate of lower-level data. On the contrary, the indirect approach adjusts the lowest level of data and aggregates the time series after the adjustment process. 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 could reinforce the seasonal pattern while allowing the cancellation of some noise in the time series. UNIDO Statistics has adopted the indirect approach for country group aggregates to preserve the additive relationship between data. Aggregates published by NSOs, with a presence of seasonality, are mostly seasonally adjusted using the direct approach. For other aggregates on the country level created by UNIDO Statistics, a case-by-case study is used to determine the proper approach.

In seasonal adjustment methods, benchmarking entails a procedure in which the annual sums of seasonally adjusted data are rendered equal to the annual sums of the non-seasonally adjusted data. Even though this ensures consistency between seasonally adjusted and raw data over the year, UNIDO Statistics does not implement this procedure, as it may degrade the quality of the seasonal adjustment and thus produce non-optimal results. Hence, careful consideration is needed when using annual IIPs derived from quarterly IIPs with filtered seasonality.

3.3 Data handling

Seasonal adjustment at country level is carried out by UNIDO Statistics only in case original data with a clear presence of seasonality are available from national sources. If no seasonality is present, the original data will be taken as seasonally adjusted data. Occasionally, seasonal adjustment will be conducted in other special cases, such as now-casting. Measures to normalize the data when the IRIIP standards [10] are not followed are taken, if necessary, while preserving the underlying message of nationally published production index figures. For example, if countries publish data based on a different or outdated classification scheme.

The data handling process (see figure 4) starts with the collection of monthly or quarterly IIP data from national statistical sources as seasonally adjusted as well as unadjusted data. If only unadjusted data with a presence of seasonality is available, seasonal adjustment is carried out with the JDemetra+ software. The data is further processed, i.e. chain-linked or rebased to the current base year 2015 and stored in an appropriate database as a time series preferably from 2005 onwards. Depending on the different types of publication, the data is published as monthly and/or quarterly seasonally adjusted data as well as unadjusted data.

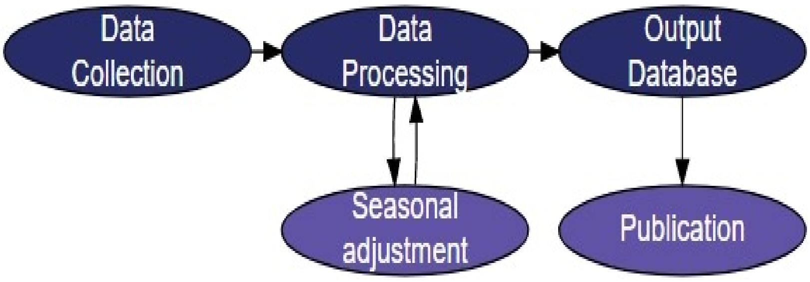

Figure 4: Data handling process in UNIDO Statistics

JDemetra+ requires at least three years of data to compute a seasonal adjusted time series and often more than four years of data to obtain an adequate adjustment and a more reliable forecast. Therefore, short time series or insufficient observations causes problems with the performance of seasonal adjustments and forecasts.

Within JDemetra+, the user has to choose between two options: TRAMO/SEATS and X13. As discussed earlier, UNIDO Statistics uses TRAMO/SEATS. The details of each available specification for TRAMO/SEATS are described in table 2, where the RSAfull specification is the most comprehensive and the default for JDemetra+. With this option, the algorithm selects the best model, which accelerates the process by reducing manual intervention by the user. The results using RSAfull are often very reliable, but modification by the user is still possible.

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 program 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 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).
RSAfull	As RSA5, but the working day specification uses 1 parameter.

Table 2: Available specification options in TRAMO/SEATS

As the seasonal adjustment process is very time consuming, UNIDO Statistics uses the RSA-full specification, whenever possible. In various cases, such as undetected seasonal outliers or a bad quality of the results, modifications have to be made manually; the major parameters for the modifications can be found in table 3.

Reason	Modification	Description
Unstable time series (e.g. possible changes of methods)	Shortening the time span	To correct annual totals, autocorrelation and normality problems as well as ”zero” observations.
Residual seasonality	Imposing the default (Airline) model (0,1,1)(0,1,1)) (check ”Automatic” in ”ARIMA”)	In rare cases, the ARIMA model defined by TRAMO cannot capture seasonality. The automatically selected ARIMA model may produce irregular components correlated with seasonal components. The Airline model is a good benchmark model in such cases and its flexibility may offer a solution.
Undetected trading days effects	Imposing trading days effect (change in ”REGRESSION” ”Automatic”)	Although the software pre-tests the presence of the trading days effect, it sometimes does not capture almost significant effects. The residual trading day graph can help, if the trading days effects should be imposed manually or not.
Definition error in main results	No trading day effect (change in ”REGRESSION” ”Option”)	If the software’s main results indicate ”definition” or ”seas-irr correlation” error, the trading days effect should be removed from the model.
Seasonal outliers	Include seasonal outliers (”REGRESSION” Pre-specified outliers)	With seasonality at certain time intervals but not for the entire time series, the automatically identified model fails to capture seasonality. The model should be identified by the expert using seasonal outliers or intervention variables.
Undetected significant outliers	Decreasing the critical value (change in ”OUTLIERS” ”Use default critical value”)	JDemetra+ defines a default critical value to detect outliers. The expert can change the value manually, if significant outliers are evident in the graphs.

Table 3: Explanation of manual modifications for seasonal adjustment

3.4 Examples

The seasonally adjusted and original IIP data for China and the Russian Federation are illustrated in figure 5 and figure 6. It is easy to identify the seasonal pattern of the original time series in both figures. The production index drops gradually by the end of the fourth quarter and rises strongly in the first quarter in almost every year, which could be attributed to variations in demand, changes in working hours during the specific period 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. For instance, the economic disruptions at the beginning of 2020 due to COVID-19 remain in both seasonally adjusted time series. The previous financial crisis of 2008/09 is also visible in the Russian example, whereas Chinese production was not severely affected. Other random disruptions and unusual movements understandable in economic terms (e.g. the consequences of economic policy, large scale orders or strikes) will also stay visible.

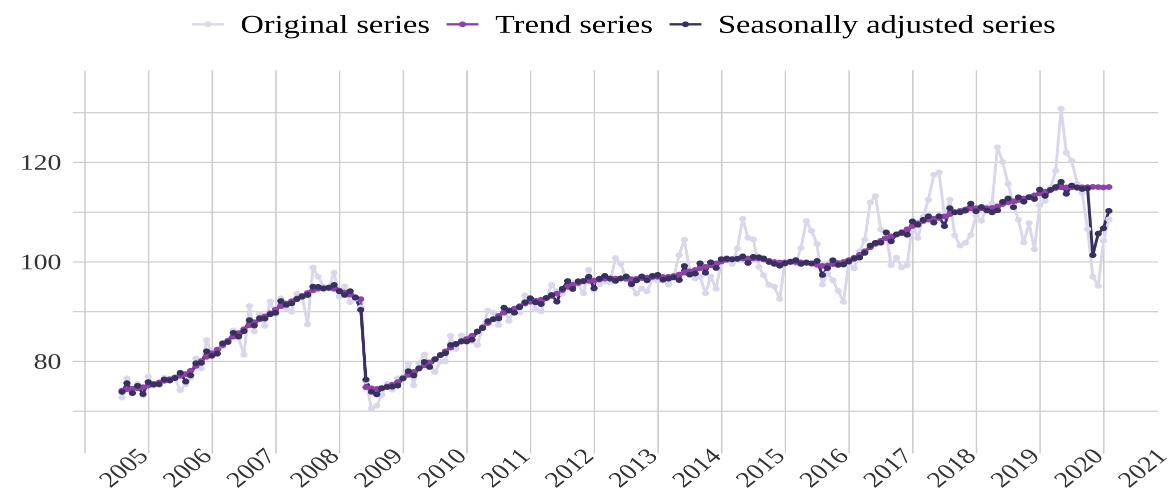

Figure 5: Seasonally adjusted and original IIP for the Russian Federation

The differences in growth rate figures before and after performing seasonal adjustment are shown in table 4. This provides a better picture of why seasonal adjustment is necessary and how it helps to reveal the real movements and turning points in manufacturing production, which may be impossible or difficult to identify due to seasonal movements. For instance, in the second quarter of 2019, the original data for China indicated a decline of 2.4 percent in manufacturing production compared to the previous quarter. Seasonally adjusted data,

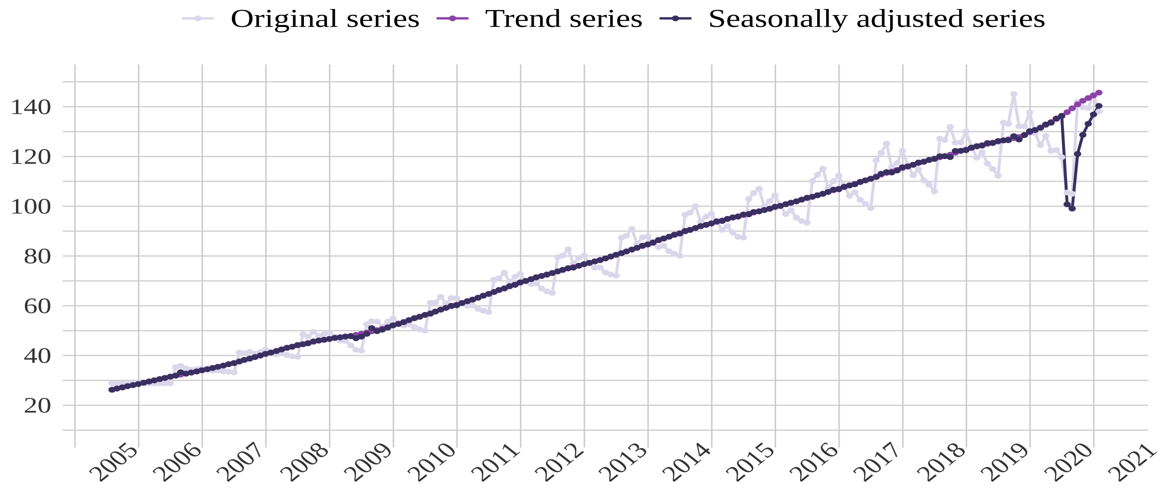

Figure 6: Seasonally adjusted and original IIP for China

	QIV 2008		QII 2019		QII 2020	
Index	Russian Federation	China	Russian Federation	China	Russian Federation	China
Original	-15.06	-7.94	-0.64	-2.38	-11.85	20.39
SA	-14.15	0.13	0.88	0.5	-9.14	20.02

Table 4: IIP percentage changes compared to the previous period

on the other hand, show a growth of 0.5 percent. In the same quarter, a similar effect for the Russian Federation was visible: instead of a decline of 0.6 per cent (according to original data), an increase of 0.9 per cent was estimated.

3.5 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, attributed to additional or improved information (in terms of coverage or reliability). Second, revisions 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. Especially the in-time modelling of unpredictable and unstable events (such as the crisis due to COVID-19) are based on assumptions made on future developments, which make data revisions inevitable as soon as new data or information becomes available.

However, it needs to be considered that one additional observation could result in revisions of the entire series of seasonally adjusted data. 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.

Prior to developing a revision policy, users' needs and available resources for implementing the policy have to be considered. The policy should address the frequency and the precision of seasonally adjusted data, the frequency of publication of revised data as well as the timing of the publication of revised seasonally adjusted data and revised unadjusted data.

Seasonally adjusted data should be revised 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 unadjusted data. The public should be informed about the average revisions of important seasonally adjusted macroeconomic variables, which have been observed in the past. The revision policy should not lead to the publication of sub-optimal seasonally adjusted data, which could mislead users when conducting economic assessments.

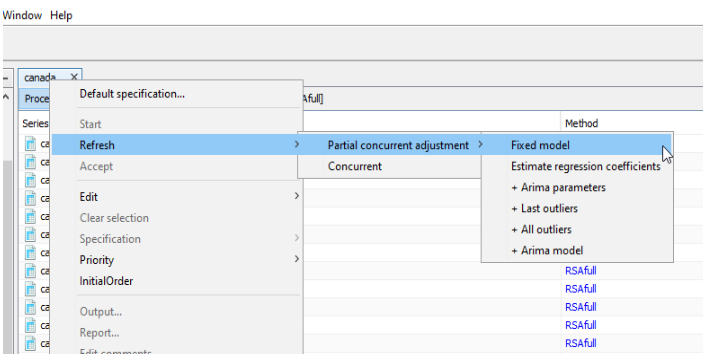

Figure 7: Revision options in JDemetra+

JDemetra+ offers three types of workspace refreshment when new data become available (see figure 7):

- Current adjustment: adjustment with fixed specification while user defined regression variables can be updated (Partial concurrent - Fixed model in JDemetra+).
- Partial concurrent adjustment: re-estimates respective parameters and factors each time new or revised observations become available (Partial concurrent - estimate regression coefficients, ARIMA parameters and others, depending on the chosen parameters in JDemetra+).
- Concurrent adjustment: adjustment performed without any fixed specifications (Concurrent in JDemetra+).

UNIDO Statistics defines the new specification and revises the models based on a partial concurrent revision with a fixed model when all sub-annual IIP data become available, mostly when the last period of the given calendar year is released. This type of 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.

When revisions affect two or more years, the specification, such as outliers and regression parameters, have to be re-identified and re-estimated. It is usually acceptable to revise the last three to four years of seasonally adjusted data before the beginning of the revision period of the unadjusted data, while earlier data should remain unchanged.

3.6 Data quality and costs

Seasonal adjustment is a time-consuming procedure, demanding significant computer and human resources. More than 800 time series of IIPs of around 50 countries at the sector level are seasonally adjusted by UNIDO Statistics for every quarterly report. According to preliminary results, more than 90 percent of the available unadjusted time series had significant underlying seasonality and/or trading days effects.

UNIDO Statistics performs 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. It is recommended that the method and software used are explicitly mentioned in the metadata accompanying the series. Unsuitable or low-quality seasonal adjustment can generate misleading results and increase the probability of false signals. The presence of residual seasonality as well as over-smoothing, are concrete risks that can negatively affect the interpretation of seasonally adjusted data.

The quality of the original data plays a key role for reliable and meaningful seasonally adjusted data. For instance, series with too many zero values will cause estimation problems for JDemetra+ and the software may estimate negative values for those units or may simply send an error message.

Possible outlier values (see chapter 2.1.4) in series have to be identified and considered in the model to prevent future problems. Furthermore, changes in the time series pattern might lead to modifying the time span of data to be used for seasonal adjustment, as 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 as well as of year-by-year growth rates are useful for this purpose.

It is important to study the data and visualize the original time series before running JDemetra+ for seasonal adjustment. Graphic illustration of the time series is a powerful tool to identify quality issues.

3.7 Data presentation

The most important UNIDO publication referring to IIPs is the quarterly report of World Manufacturing Production, whose main objective is to provide an overview of the current growth trends of world manufacturing production by country groups and major industries. Therefore, the report contains two sets of growth indicators for world manufacturing output:

- as compared to the previous quarter, or
- as compared to the same period of the previous year.

These growth figures refer to the change in MVA or more precisely, to the change in production (as measured by IIP). For quarterly reports, the growth figures are transcribed to the growth percentage of a quarter-on-quarter comparison:

$$G_q^{QoQ} = \left(\frac{I_q}{I_{q-1}} - 1\right) \times 100$$

or a year-on-year comparison:

$$G_q^{YoY} = \left(\frac{I_q}{I_{q-4}} - 1\right) \times 100.$$

For monthly IIPs, the growth figures are transcribed to the growth percentage of a month-on-month comparison:

$$G_m^{MoM} = \left(\frac{I_m}{I_{m-1}} - 1\right) \times 100$$

or a year-on-year comparison:

$$G_m^{YoY} = \left(\frac{I_m}{I_{m-12}} - 1\right) \times 100.$$

While the first set of growth indices represents more recent growth trends and allow to study short-term developments, the second set provides more stable estimates when analyzing a country's manufacturing performance, as year-on-year comparisons help mitigate many undetected seasonal or calendar variations. The quarterly report uses year-by-year comparisons instead of quarter-on-quarter comparisons, as national data are not always seasonally adjusted and there is thus a chance that certain hidden seasonal or calendar patterns caused by a country's specific national holidays or other particularities might remain in the adjusted series when carrying out seasonal adjustment at the international level. A similar reasoning is used for the monthly report, where year-on-year comparisons are preferred to month-over-month comparisons.

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 and calendar effects (e.g. moving holidays).

Therefore, countries with no seasonal adjustment experience are encouraged to compile, maintain and update their national calendars or, as an alternative, to supply a list of past and future public holidays including information on compensation holidays.

4 Conclusion

A decision regarding a specific seasonal adjustment approach for each time series under review needs to be made, and once a method has been selected, it is not advisable to change it too frequently. This would lead to unnecessary revisions that could affect data users. The same holds for the revision policy. Before initiating a seasonal adjustment process, 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, the seasonally adjusted time series should be made available.

One common misconception is that seasonal adjustment hides outliers. Unusual or unpredictable events, such as the impacts of the COVID-19-pandemic in 2020, play an important role for modelling and analyzing economic data. The seasonal adjustment needs to be conducted carefully with special focus on potential outliers. Seasonally adjusted data referring to uncertain times might face bigger data revisions as new information can lead to rejecting or adapting assumptions made in previous periods. New information must always be evaluated carefully and considered in the model.

Seasonally adjusted data provide more readily interpretable measures of changes that occur in a given period and reflect real economic movements without misleading seasonal patterns. However, users of seasonally adjusted data should be aware that the data's suitability for econometric modelling purposes needs to be carefully considered.

References

[1] George E.P. Box and Gwilym M. Jenkins. *Time Series Analysis, Forecasting and Control.* HoldenDay, 1970.

[2] Chung Chen and Lon-Mu Liu. "Joint Estimation of Model Parameters and Outlier Effects in Time Series". In: *Journal of the American Statistical Association* 88 (1993), pp. 284–297.

[3] Eurostat, European Commission. *Handbook on Seasonal Adjustment 2018 Edition.* 2018. ISBN: ISBN 978-92-79-80170-9. URL: `https://ec.europa.eu/eurostat/documents/3859598/8939616/KS-GQ-18-001-EN-N.pdf`. (accessed: October 5, 2020).

[4] Víctor Gómez and Agustín Maravall. *Programs TRAMO and SEATS Instructions for the user.* 1996. URL: `https://www.bde.es/f/webbde/SES/Secciones/Publicaciones/PublicacionesSeriadas/DocumentosTrabajo/96/Fich/dt9628e.pdf`. (accessed: October 16, 2020).

[5] Sylwia Grudkowska. *JDemetra+ User Guide.* 2015. URL: `https://ec.europa.eu/eurostat/cros/system/files/jdemetra_user_guide.pdf`. (accessed: October 5, 2020).

[6] Sylwia Grudkowska. *JDemetra+ Reference Manual Version 2.1.* 2016. URL: `https://ec.europa.eu/eurostat/cros/system/files/jdemetra_reference_manual_version_2.1_0.pdf`. (accessed: October 9, 2020).

[7] Agustín Maravall, Roberto López Pavón, and Domingo Pérez Cañete. *Reliability of the Tramo-Seats Automatic Identification of Regression-Arima Models.* 2013. URL: `https://www.researchgate.net/publication/310460158_RELIABILITY_OF_THE_AUTOMATIC_IDENTIFICATION_OF_ARIMA_MODELS_IN_PROGRAM_TRAMO_External_collaboration_with_the_Bank_of_Spain`. (accessed: October 5, 2020).

[8] Ruey Tsay. "Time Series Model Specification in the Presence of Outliers". In: *Journal of the American Statistical Association* 81 (1986), pp. 132–141.

[9] UNIDO Statistics. *World Manufacturing Production (Methodological Note).* 2020. URL: `https://stat.unido.org/content/publications/world-manufacturing-production-%2528methodological-note%2529`. (accessed: November 5, 2020).

[10] Department of Economic UNSD United Nations and Statistics Division Social Affairs. *IRRIP: International Recommendations for the Index of Industrial Production 2010.* 2010. URL: `https://unstats.un.org/unsd/industry/Docs/F107_edited.pdf`. (accessed: October 5, 2020).

[11] Center for Statistical Research U.S. Census Bureau and Methodology. *X-13ARIMA-SEATS Reference Manual.* 2017. URL: `https://www.census.gov/ts/x13as/docX13AS.pdf`. (accessed: October 9, 2020).