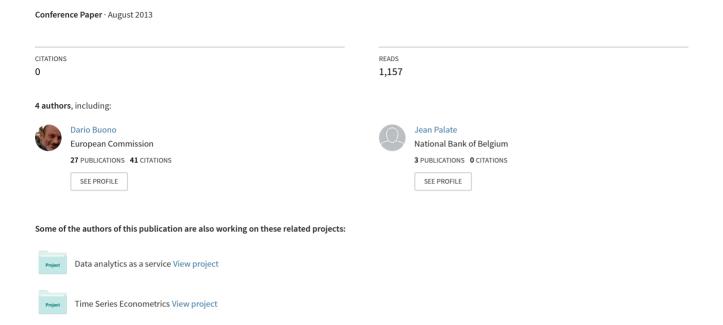
Advanced Tools for Time Series Analysis and Seasonal Adjustment in the New JDemetra+



Advanced Tools for Time Series Analysis and Seasonal Adjustment in the New JDemetra+

Sylwia Grudkowska¹, Dario Buono², Jean Palate³, Wojciech Ciebiera⁴

¹Narodowy Bank Polski, Swietokrzyska 11/21, 00-919 Warsaw, Poland¹

²European Commission, Eurostat, Unit C0: "Macroeconomic Imbalances Procedure"

³National Bank of Belgium, Boulevard de Berlaimont 14, 1000 Brussels, Belgium

⁴Narodowy Bank Polski, Swietokrzyska 11/21, 00-919 Warsaw, Poland

Abstract

JDemetra+ is a new open source tool for seasonal adjustment (SA) that enables the implementation of the ESS Guidelines on SA. It has been developed to provide a set of reusable and extensible components, following a standard technology, compatible with the environment of most European statistical institutions. JDemetra+ is not only a user-friendly graphical interface, comparable to its predecessor, Demetra+, but also a set of open Java libraries that can be used to deal with time series related issues like the SA processing of large-scale data sets, non-standard SA methods, the development of advanced research modules, temporal disaggregation, benchmarking and business cycle analysis. JDemetra+ is built around the concepts and the algorithms used in the two leading SA methods, i.e. TRAMO/SEATS and X-12-ARIMA/X13-ARIMA-SEATS. They have been reengineered, following an object-oriented approach, which allows easier handling, extensions or modifications. The paper aims to describe the architecture of JDemetra+ and to present its main functionalities: benchmarking, temporal disaggregation, seasonality tests and a tool for comparing the direct vs. the indirect approach.

Key Words: Seasonal adjustment, Statistical tools, Time series, X-13-Arima-Seats, Tramo-Seats

1. Introduction

This paper discusses some recent developments in the area of seasonal adjustment software. To familiarize the readers with the subject, the historic background has been outlined below.

Seasonal adjustment (SA) is a widely used technique for estimating and removing seasonal and calendar-related movements from time series in order to provide the users with data allowing clear understanding of economic phenomena. There are several efficient seasonal adjustment methods to produce reliable results for large data sets. The most widespread and recommended are X-12-ARIMA ² (X-13ARIMA-SEATS ³)

¹ The views expressed in the paper are those of the authors and not necessarily those of the affiliated institutions.

² X-12-ARIMA is a seasonal adjustment program developed and supported by the U.S. Census Bureau. It includes all the capabilities of the X-11 program, which estimates trend and seasonal component using moving averages. X-12-ARIMA offers useful enhancements including:

developed at the U.S. Census Bureau and TRAMO/SEATS⁴ developed by Victor Gómez and Agustín Maravall, from the Bank of Spain. Both original seasonal adjustment programs are written in FORTRAN, which is currently perceived obsolete. As FORTRAN is not a strictly object-oriented⁵ language, the SA algorithms written in it might be applied to solve time series related issues, but generally they are not designed for reusability. The limitations of this declining language make their maintenance increasingly difficult.

Under these circumstances, the Eurostat⁶, which has been playing a leading role in promotion, development and maintenance of publically available software solution for SA, developed in 2002 a new software, called Demetra, to ease an access of non-specialists to the SA methods. Demetra offered a user-friendly interface to the two SA algorithms: TRAMO/SEATS and X-12-ARIMA and aided the comparison of the output from those two algorithms. Regrettably, because of lack of sufficient product development and handling of errors, the initially successful program became outdated soon.

In 2009, the European Statistical System launched its Guidelines on Seasonal Adjustment⁷. It immediately became quite obvious that Demetra could not be adapted to

extension of the time series with forecasts and backcasts from ARIMA models prior to seasonal adjustment, adjustment for effects estimated with user-defined regressors, additional seasonal and trend filter options, alternative seasonal-trend-irregular decomposition, additional diagnostics of the quality and stability of the adjustments, extensive time series modelling and model selection capabilities for linear regression models with ARIMA errors. For basic information on the X-12-ARIMA program see 'X-12-ARIMA Reference Manual' (2007). More information on X-12-ARIMA can be found at http://www.census.gov.

³ X-13ARIMA-SEATS is a seasonal adjustment program developed and supported by the U.S. Census Bureau that contains two seasonal adjustment modules: the enhanced X-11 seasonal adjustment procedure and ARIMA model based seasonal adjustment procedure from the SEATS seasonal adjustment program developed by Victor Gomez and Agustin Maravall (2013). For basic information on the X-3ARIMA-SEATS program see 'X-13ARIMA-SEATS Reference Manual' (2013). More information on X-12-ARIMA can be found at http://www.census.gov.

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

⁵ Object-oriented language is a programming language in which it is possible to define the data type of a data structure and types of operations, called functions. Those operations are then applied to the data structure, so it can be viewed as an object that includes both data and functions. Due to this solution, the object-oriented languages, in contrast to the procedural programming techniques, enable programmer to develop modules that do not need to be modified once a new type of object is added and to create objects that inherits given features from the already defined objects. Among the most commonly used object-oriented languages are Java and C++.

⁶ Eurostat is the statistical office of the European Union. Its task is to provide the European Union with statistics at European level that enable comparisons between countries and regions. More information at http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/.

⁷ Endorsed by the Statistical Programme Committee, the European Statistical System (ESS) Guidelines on Seasonal Adjustment aim to harmonize European practices and to improve the

new requirements implied by the Guidelines. Eurostat, in cooperation with the National Bank of Belgium (NBB), started a project aiming to develop an improved version of Demetra. The new software, called Demetra+⁸, was released in 2012. This tool provides a common approach for seasonal adjustment using TRAMO/SEATS and X-12-ARIMA methods, which is more coherent with the Guidelines. It includes unified graphical interface and input/output diagnostics for the two methods. From an IT perspective Demetra+ consists in a .NET encapsulation of the two original FORTRAN modules. Therefore it cannot be used under IT environments other than Windows.

Although Demetra+, which is used for production of SA time series in many European institutions, including ISTAT⁹, INSEE¹⁰, Eurostat, Turkstat¹¹, NBP¹², proved to be successful within the SA community, it seems that in long-term perspective it will not meet the growing users' needs. The main reasons for this are the IT constraints implied by the use of the FORTRAN components, which act as black-boxes. More especially, the current solution is not adapted to extensions or to reuse and its maintenance is cumbersome. Basing on such experience Eurostat took an initiative to create a platform independent, extensible seasonal adjustment software. The task, undertaken by NBB, resulted in launching JDemetra+ (2012). Although numerous improvements to the software has been implemented, as well in terms of layout as in terms of functionalities, there is little difference between JDemetra+ and Demetra+ from the typical user point of view. The actual added value of JDemetra+ to the development of the SA tools lies in the re-writing of the original FORTRAN codes of X-12-ARIMA and TRAMO/SEATS in JAVA, following a real object-oriented approach. This solution is a crucial factor that enables the long-term maintenance of the tool, integration of the libraries in the IT environments of many institutions and re-use of the modules and algorithms for other purposes. The structure of JDemetra+, its unique solutions and resulting capabilities are explored in the subsequent sections of this paper.

2. Object-oriented approach

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

comparability of infra-annual national statistics as well as enhance the overall quality of the European Union and the euro area aggregates. The Guidelines on Seasonal Adjustment cover all the key steps of the seasonal and calendar adjustment process. They discuss both the theoretical aspects and practical implementation of seasonal adjustment issues.

⁸ Grudkowska, S. (2013), 'Demetra+ User Manual'.

⁹ ISTAT (Italian National Institute of Statistics) is a public research organisation that is the main producer of official statistics for Italy in the service of citizens and policy-makers. More information at http://www.istat.it/en/.

¹⁰ INSEE (France's National Institute of Statistics and Economic Studies) is a Directorate General of the Ministry of the Economy, Finance, and Industry. It collects, produces, analyses and disseminates information on the French economy and society. More information at www.insee.fr.

¹¹ Turkstat (Turkish Statistical Office) is a public research organisation responsible for collecting and compiling statistics for Turkey. More information at www.turkstat.gov.tr.

¹² NBP (Narodowy Bank Polski) is the central bank of the Republic of Poland. More information at www.nbp.pl.

3. JDemetra+ structure

In fact, JDemetra+ is a set of open Java libraries. Technically, it consists of a set of pure Java libraries called JDemetra+, and of NbDemetra, which is based on the NetBeans platform¹³. The structure of JDemetra+ is presented on Figure 1. The pure Java part is presented in orange while NBDemetra components are highlighted in blue.

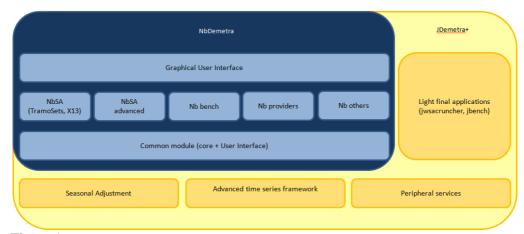


Figure 1: JDemetra + structure.

JDemetra+ libraries contain "pure" Java modules, which are independent of any framework. The most important one is the library called *jtstoolkit.jar*, situated in *Seasonal Adjustment* part, which comprises all the algorithmic parts (including TRAMO/SEATS and X-13ARIMA-SEATS). Consisting of approximately 1000 classes it is the largest module of JDemetra+. The skeleton of the SA routines reflects as much as possible the structure of the original programs. However, they rest on the generic routines provided by the independent classes for most of the computations. Therefore, the user should not expect that SA in JDemetra+ performed in a fully automatically way would produce the same output as original SA programs. The *jtstoolkit.jar* library doesn't depend on any external package. The integration of SA routines in in-house developments should be based on that module.

NbDemetra can be considered as a set of plugins for a time series analysis in general and for a seasonal adjustment in particular. It provides the graphical interface to the actual algorithmic modules. The interface enables the user to call the functionalities from the lower layer (*NbSA*, *NB advanced*, and so on) and to determine the parameters of the process. To perform the actual computations NbDemetra uses Java libraries from JDemetra+. Due to applying NetBeans framework in NbDemetra area it is possible to integrate the new features by outside teams with no or few impacts on the current application. In this way the long term maintenance of the software has been simplified. This issue is further discussed in 2.1.

¹³ NetBeans is an integrated development environment (IDE) which allows user to build rich client applications. NetBeans platform is a generic framework for swing applications which provide user with: user interface management, user settings management, storage management, windows management, wizard framework, NetBeans visual libraries and integrated development tools. NetBeans is a free open source with support for Java programming.

Common module is used to launch whole application. It enables to load the modules one by one in order to get all functionalities. This approach is provided by NetBeans framework and can be easily used by developers to make optional plugins. In such way, the user can load and install plugin which will be used in a similar way to those originally provided with NbDemetra.

Peripheral services include modules that are mainly aimed at providing I/O facilities, particularly different providers, like XML, ODBC, JDBC, SDMX, EXCEL.

JDemetra+ also includes two applications: WSACruncher and JBench, to facilitate the seasonal adjustment process. WSACruncher is a command line tool, designed to be used with JDemetra+, that re-estimates all the multi-processes defined in a workspace¹⁴. It can be launched under system command line or in the batch files. In such way the user can combine other applications like Matlab or SAS with Demetra+.

3.1. Plugins

NbDemetra is composed of a collection of plugins that can be (de-)activated following the user's needs. The current version comes with plugins for providing time series (ODBC, JDBC, Spreadsheet) and with plugins for seasonal adjustment (TRAMO/SEATS and X-13ARIMA-SEATS in the *NbSA* plugin, structural models in the *NbSA* advanced plugin). By (de-)activating plugins, the user can simplify/enrich the features displayed in the main application (menus, workspace, time series providers, tools, etc.).

NetBeans provides a graphical component for the management of plugins. The tool enables manual check if some updates of the already installed plugins are available (*Updates* panel), downloading all plugins that are related to JDemetra+ (*Available plugins* panel), installation of the new plugins from the local machine and their integration with the software (*Downloaded* panel) and adding the update centres, which are the places to hold plugins (*Settings* panel).

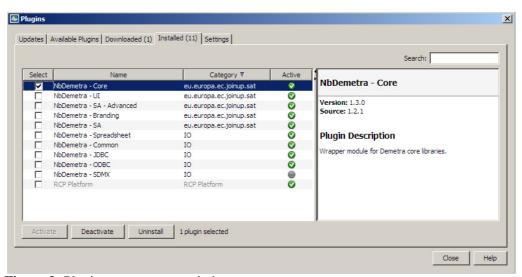


Figure 2: Plugins management window.

¹⁴ A workspace contains the complete description of all the documents (e.g. processing, specifications, calendars) saved by the user during a session of JDemetra+.

To sum up, JDemetra+ is easily expandable. Moreover, almost all new plugins or new versions of existing plugins can be downloaded from the Internet. They may modify the layout of the application, provide new tools that will be available through extended menus or even change default components (for example by adding new graphical panels to the existing ones). The exemplary improvement of the graphical interface made by the user is a report from a seasonal adjustment of large dataset designed to satisfy Narodowy Bank Polski needs in this field. In comparison with the default report, the enhanced one includes information about report's author, tool to customise the content of the report (include/exclude given part of the report), more output options (export to Excel file) and enriched information about SA process, i.e. frequency of the series.

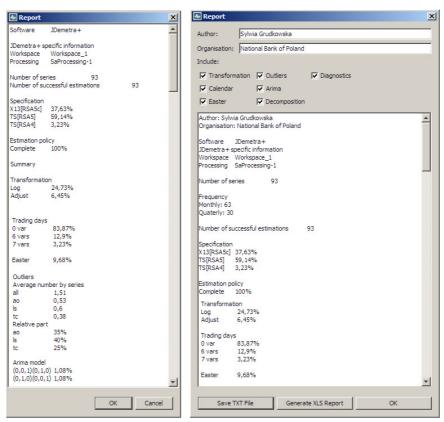


Figure 3: A comparison of default (on the left) and improved (on the right) report from seasonal adjustment.

Although plugins may replace or modify the original plugins, the majority of users prefer them to provide additional options without removing the default ones. The example below illustrate the case in which both default and enhanced reports are available.

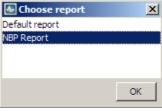


Figure 4: Choice of metadata report.

3.2. Reusability and extensibility

JDemetra allows external IT teams to create their own panels for presenting specific aspects of SA processes without modifying the core modules. In greater detail, each SA process is presented in a window that contains a navigation tree on the left, with a list of the different visual aspects available for that process (e.g. Specification, Main results); the panel corresponding to the selected item is displayed on the right side of the window.

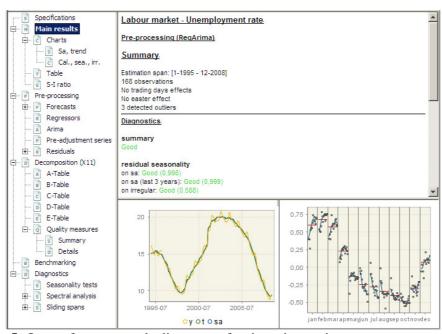


Figure 5: Output from seasonal adjustment of a given time series.

External developers can add to the navigation tree their own items and provide for them the corresponding presentation panels. In this way the users may enrich the default results by additional information. The example of a functionality that may be added here is a revision history diagnostic¹⁵.

4. Functionalities

Basically, JDemetra+ incorporates all the capabilities of the latest version of Demetra+, including:

- seasonal adjustment with TRAMO/SEATS and X-12-ARIMA of a single time series and large data sets, performed in an automatic way using pre-defined specifications as well as by setting all the parameters directly by the user;
- defining and using in SA routines different kinds of calendars (national, composite and chained) that enables to specify complex calendar variables;

¹⁵ Revision history is a tool that illustrates the changes in the seasonally adjusted series and trend-cycle component which take place as new observations are added to the time series. It has been originally developed for X-12-ARIMA. For more details see Findley, D.F., Monsell, B.C., Bell, W.R., Otto, M.C., and Chen, B-C. (1998).

- automatic detection of seasonal outliers ¹⁶;
- user-defined regression variables that can be integrated with the model for pretreatment¹⁷;
- various ways to display and analyse the data (e.g. periodogram, Tuckey spectrum, ACF function);
- functionalities to facilitate the regular data production, like define, save and refresh the seasonal adjustment process using several revision policies.

Except from TRAMO/SEATS and X-12-ARIMA, JDemetra includes also X13-ARIMA-SEATS, structural models¹⁸, Mixed Airline models¹⁹ and generalized Airline models²⁰. In comparison with Demetra+, the set of spectral analysis tools in JDemetra+ has been enriched by Tuckey Spectrum.

Although at present JDemetra+ concentrates on seasonal adjustment, it can be also used for other purposes, like time series modelling and forecasting using TRAMO and RegARIMA models. These models can be identified manually or automatically by means of pre-defined specifications. The output for this functionality presents the ARIMA model structure, including regression variables, forecasts with forecasts' errors and analysis of residuals.

In order to fulfill ESS Guidelines requirements, JDemetra+ incorporates a module for comparison of the results from direct and indirect seasonal adjustment ²¹. The tool performs several seasonality tests for both types of seasonal adjustment and analysis differences between direct and indirect seasonal adjustment. The typical view displayed for this functionality is presented below.

¹⁶ Seasonal outlier is a level shift that only operates on a particular period (e.g. month of December, fourth quarter of the year). For more information see Kaiser, R., and Maravall, A. (1999).

¹⁷ Pre-treatment is an adjustment of a time series for variations caused e.g. by calendar effects and outliers.

¹⁸ Structural time series models are regression models in which explanatory variables are functions of time and the parameters change with time. The distinctive feature of the structural models is their ability to present the changes in behaviour of a given time series by taking its different components, like trend-cycle, seasonality and irregularity, as stochastic processes governed by random disturbances.

¹⁹ Trimbur, T.M., and Bell, W.R. (2008).

²⁰ Aston, J.A., Findley, D.F., Wills, K.C., and Martin, D.E.K. (2004).

²¹ Direct seasonal adjustment means that all time series, including aggregates, are seasonally adjusted individually. On the contrary, indirect seasonal adjustment takes place if the seasonally adjusted figures for the time series are obtained by combining the estimates for two or more directly adjusted series. The problem of direct/indirect SA is discussed in ESS Guidelines on Seasonal Adjustment (2007).

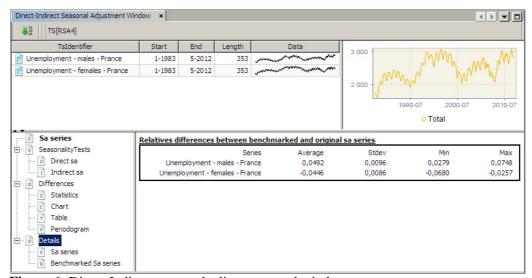


Figure 6: Direct-Indirect seasonal adjustment tool window.

The next improvement in JDemetra+, in comparison with Demetra+, is a seasonality test tool. It enables to test the time series for presence of seasonality before the actual seasonal adjustment is performed. It provides the user with knowledge about characteristic of seasonal movements in the time series. The tool includes six tests:

- Test on autocorrelation on seasonal lags, i.e. Ljung-Box test that checks if significant autocorrelations at the seasonal lags exist;
- Friedman test, which is a non-parametric method for testing that samples are drawn from the same population or from populations with equal medians by checking the significance of the month (or quarter) effect in the regression equation;
- Kruscall-Wallis test, which is a non-parametric test used for comparing samples from two
 or more groups;
- Spectral peaks, which considers existence and significance of seasonal peaks in a Tuckey periodogram and in an auto-regressive spectrum;
- Tuckey periodogram, which checks the sum of the values of a Tuckey periodogram at seasonal frequencies;
- Seasonal dummies test which checks the significance of fixed seasonal dummies in the regression model.

The exemplary output is presented in Figure 7.



Figure 7: Example of seasonality tests output.

4.1. Benchmarking

JDemetra+ provides the benchmarking²² of uni-variate series, based on the method of Cholette, which generalizes, amongst others, the additive and the multiplicative Denton's procedure. That feature is integrated in the SA algorithms²³. The typical output from benchmarking analysis is presented on figure 8.

_

²² In seasonal adjustment methods benchmarking means the procedure that ensure the consistency over the year between adjusted and not seasonally adjusted data. More generally, benchmarking consists in adjusting a given series to a target, following a method that preserves as much as possible the original series (based on some well-defined criterion).

possible the original series (based on some well-defined criterion).

23 Usually, benchmarking is not a recommended practice. So, by default, it will be disabled. However, in some context (for instance in the Quarterly National Accounts), users might require annual consistency.

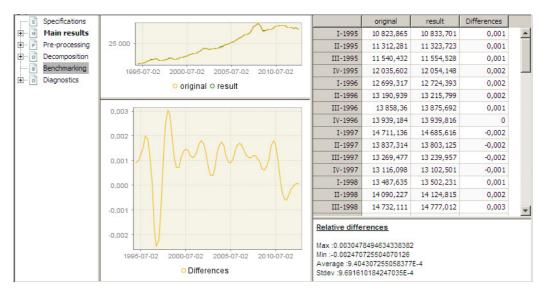


Figure 8: Benchmarking tool window.

In the case of direct seasonal adjustment of aggregated series (see above), users could also want to reconcile the seasonally adjusted sum with its seasonally adjusted components. The Cholette's method has been extended to the multi-variate case (with or without temporal constraints).

A separate plug-in, called *NbDemetra-Benchmarking*, which is not part of the core of NbDemetra, provides a graphical interface on the multi-variate benchmarking outside the context of seasonal adjustment. It allows benchmarking small to medium sized sets of data following temporal and/or contemporaneous constraints. The last ones may represent, for instance, the margins of multi-ways tables or hierarchical structures.

4.2. Calendars

Demetra+ offers a wide variety of calendar variables, including trading and working days effects, leap year effect, Easter and Easter related holidays. In contrast to Demetra+, JDemetra+ supports both stock and flow trading days effects. The calendars are corrected for long-term effects. This correction take into account the distribution of the Easter days.

4.2.1. Weighted calendars

JDemetra+ supports weighted holidays which allow to specify a given day as partially working and partially non-working day. This option facilitates developing of calendar variables that describe an actual economic activity in a reliable way.

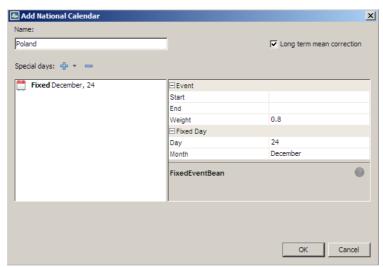


Figure 9: Window to define a national calendar.

4.2.2. Reduction of the number of calendar variables

Instead of considering the reduction of the number of calendar variables, JDemetra+provides a solution that chooses automatically the best number of regression variables. That option is available for TRAMO only²⁴. The choice is based on F-tests (Wald tests) computed on the nested models without trading days, with working days (2 variables) and with full trading days (7 variables).

For each pre-treatment model the user may also specify the number of calendar variables by hand.

5. Summary

JDemetra+ is the newest member of Demetra family, designated not only for pure seasonal adjustment purposes but also to the other time series related issues. In contrast to its predecessors, it is a reusable and extensible tool that follows an object-oriented approach. This feature enables numerous concepts of JDemetra+ to be reused in other projects, e.g.: JEcotrim ²⁵ (for temporal disaggregation and benchmarking), outliers detection, estimation of missing values and forecasting. The major advantage of JDemetra+ is possibility for outside programmers to develop and integrate with JDemetra+ additional modules, which is a unique solution among seasonal adjustment software. The modification and extension of the code can be realised by e.g.: modification of existing functionalities; adding new data providers due to genetic serialization functionality; implementing additional diagnostics and output; integrating new seasonal adjustment methods (using batch processing).

JDemetra+ is a finished product. However it is quite certain that the need for new functionalities and improvements will appear in the future. Therefore the basic

²⁴ The solution will be a part of the future version of the reference TRAMO program. It has been chosen to not implement it for X13, to avoid differences with the reference X-13ARIMA-SEATS software.

²⁵ JEcotrim is a user-friendly and reasonably powerful tool for the time series analysis, that enables for temporal disaggregation, benchmarking, reconciliation of low frequency series and matrix balancing via complex mathematical and statistical methods.

requirement for JDemetra+ was to deliver a product that can be relatively easy to modify. The solutions implemented in JDemetra+, like usage of the NetBeans technology, enhance the improvement of the product in the future, facilitate the understanding of the code save time and resources allocated to the activities connected with the project. All these factors will make the long-term maintenance of the tool easier.

References

Aston, J.A., Findley, D.F., Wills, K.C., and Martin, D.E.K. (2004), 'Generalizations of the Box-Jenkins' Airline Model with Frequency-Specific Seasonal Coefficients and a Generalization of Akaike's MAIC'. Proceedings of the 2004 NBER/NSF Time Series Conference.

Caporello, G., and Maravall, A. (2004), 'Program TSW. Revised Reference Manual', Banco de España, http://www.bde.es/servicio/software/tramo/tswrm.pdf).

Findley, D.F., Monsell, B.C., Bell, W.R., Otto, M.C., and CHEN, B-C. (1998), 'New Capabilities and Methods of the X-12-ARIMA Seasonal-Adjustment Program', *Journal of Business and Economic Statistics*, 2, 16, 127-152.

(2007), 'ESS Guidelines on Seasonal Adjustment', http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-RA-09-006/EN/KS-RA-09-006-EN.PDF.

Grudkowska, S. (2013), 'Demetra+ User Manual', http://www.cros-portal.eu/sites/default/files//Demetra%2B%20User%20Manual%20November%202012.p df.

Kaiser, R., and Maravall, A. (1999), 'Seasonal Outliers in Time Series'. Special issue on Time Series, *Estadística, Journal of the Inter-American Statistical Institute*, 15, 101-142, 2003.

Trimbur, T.M., and Bell, W.R. (2008), 'Seasonal Heteroskedasticity in Time Series Data: Modeling, Estimation, and Testing', US Census Bureau, Research Reports Series (Statistics #2008-11)

(2007), 'X-12-ARIMA Reference Manual'. Time Series', Staff Statistical Research Division, US Bureau of the Census, http://www.census.gov/srd/www/x12a/x12down_pc.html.

(2013), 'X-13ARIMA-SEATS Reference Manual Accessible HTML Output Version'. Time Series', Staff Statistical Research Division, US Bureau of the Census, http://www.census.gov/ts/x13as/docX13ASHTML.pdf.