SPONSOR

CONTRACTOR

Statistics ♦ Economics ♦ Finance

EUROPEAN COMMISSION – EUROSTAT UNIT C2

CONTRACT: 20102.2010.003-2010.520

CALL FOR TENDER: 2010/S 109-164439

NATIONAL ACCOUNTS METHODOLOGICAL AND PRACTICAL IMPROVEMENTS: NAMP C4 (LOT 10)

IMPROVED SPECIFIC IT TOOLS FOR THE COMPILATION OF NATIONAL ACCOUNTS

JECOTRIM USER MANUAL

Project Manager, Eurostat: Luis Biedma (luis.biedma@ec.europa.eu)

Project Manager, Hendyplan:
François Libeau (francois.libeau@hendyplan.com)

UPDATED ON: 22TH MARCH 2013

TABLE OF CONTENTS

TABLE OF FIGURES

Bibliography

Barcellan, R. and Buono D. (2002). ECOTRIM Interface (Version 1.01) User Manual Draft -Including Case Studies and Examples-.

Cholette, P. (1984). Adjusting sub-annual series to yearly benchmarks. *Surv. Methodol.*, 10, 35–49.

Chow G. and Lin A.L. (1971). Best linear unbiased Interpolation, Distribution and Extrapolation of Time Series by Related Series, *The Review of Economics and Statistics* 53 (4), pp. 372-375.

Denton, F.T. (1971). Adjustment of monthly or quarterly series to annual Totals: An Approach based on quadratic Minimization, *Journal of the American Statistical Association* 66 (333), pp. 99-102.

Di Fonzo, T. and Marini, M. (2011). Simultaneous and two-step reconciliation of systems of time series: methodological and practical issues. *Journal of the Royal Statistical Society*, *Appl. Statist* 60, Part 2, pp. 143-164.

Fernandez, R.B. (1981). A methodological Note on the Estimation of Time Series, *The Review of Economics and Statistics* 63(3), pp- 471-476.

Lahr, M.L. and De Mesnard, L. (2004). Biproportional Techniques in Input-Output Analysis: Table Updating and Structural Analysis. *Economic Systems Research 16 (2), pp. 115-134.*

Litterman R.B. (1983). A Random Walk, Markov Model for the distribution of Time Series, *Journal of Business and Statistics,* 1,169-173*.*

Quenneville, B. and Rancourt, E. (2005). Simple methods to restore the additivity of a system of time series. *Workshop Frontiers in Benchmarking Techniques and Their Application to Official Statistics, Luxembourg, Apr. 7th–8th*.

1. *INTRODUCTION*

This manual is about using the software package called JEcotrim, to supply various tasks required by the Eurostat. The data for all of the examples used herein are available in a separate library of files.

The chapters are arranged in the following order: Chapter 1 introduces the general features of the software and shows the basics. Chapter 2 describes the functionalities on the main tool bar menu. Chapter 3 provides a brief yet detailed description of the different techniques covered in the package and gives the user the instructions how to reproduce those in the interactive mode. Chapter 4 is providing some recommendations for Ecotrim for Windows users.

Finally, the Annexes contain useful information on terminology and definitions. Please note that this user guide contains information about how to use the JEcotrim interface and summaries of the methods about the methodology. For more details about them, please refer to the bibliographical references.

This documentation is also focusing on the specificities of JEcotrim. The interactive interface is built on top of the JDemetra+ system, the functionalities of which are assumed to be known. Please refer to the general JDemetra+ documentation for details about them.

This guide covers the interactive mode, meaning the usage of the graphical User Interface. It allows the computation of one method at a time. For the computation of multiple steps at a time, please refer to the BATCH user guide as a separate document.

2. Common features

This chapter introduces some of the basic features of the JDemetra+ interface and of the JEcotrim module. It shows the process of installing the JEcotrim module, how to surf around windows and deployable menus and finally how to import data and workspaces.

2.1. *WHAT IS JECOTRIM?*

JEcotrim can be defined simply as an update of Ecotrim v.1.01 developed for Windows and written by 2SDA (Statistical Studies and Data Analysis, Luxembourg) on behalf of Directorate B Economic Statistics and Economic and Monetary Convergence of Eurostat, the Statistical Office of the European Union. This new software comprises a lot of remarkable improvements and corrects important errors present in the Ecotrim v.1.01.

JEcotrim is easy to use and reasonably powerful, containing procedures based on temporal disaggregation, benchmarking, reconciliation of low frequency series and matrix balancing via complex mathematical and statistical methods. As mentioned, JEcotrim takes the form of a user-friendly interface allowing the management of the program facilities.

2.1.1. JECOTRIM BASICS

There are several different ways to work in JEcotrim. The easiest and most intuitive way to use it is through its interactive mode. This mode should be familiar to most users that are accustomed to using Ecotrim v.1.01. It consists basically of open windows and dialog boxes. Thus, the user selects the series object of analysis, fills in the desired menus and corresponding specifications and clicks on the "Apply" button and the program generates the desired results together with a set of different statistics assessing the goodness and accuracy of the outputs.

Also, JEcotrim provides a set of complementary useful concepts as Ecotrim v.1.01, which depends upon the employed technique[1], such as reliability indicators, confidence intervals and growth rates.

Figure 1 shows the interface of the interactive mode in JEcotrim while

[1] Note that the aforementioned concepts are not available for mathematical methods but only for statistical ones.

Figure 2 shows the batch mode. In the former, we can find the “Providers Window” (Figure 3) and the “Workspace window” (Figure 4).

As for the other way, the batch mode, JEcotrim offers a way in which the user can load any xml file and execute it. A keyboard shortcut (Ctrl+B) is also valid to activate this batch mode.

Figure 1. Interactive mode

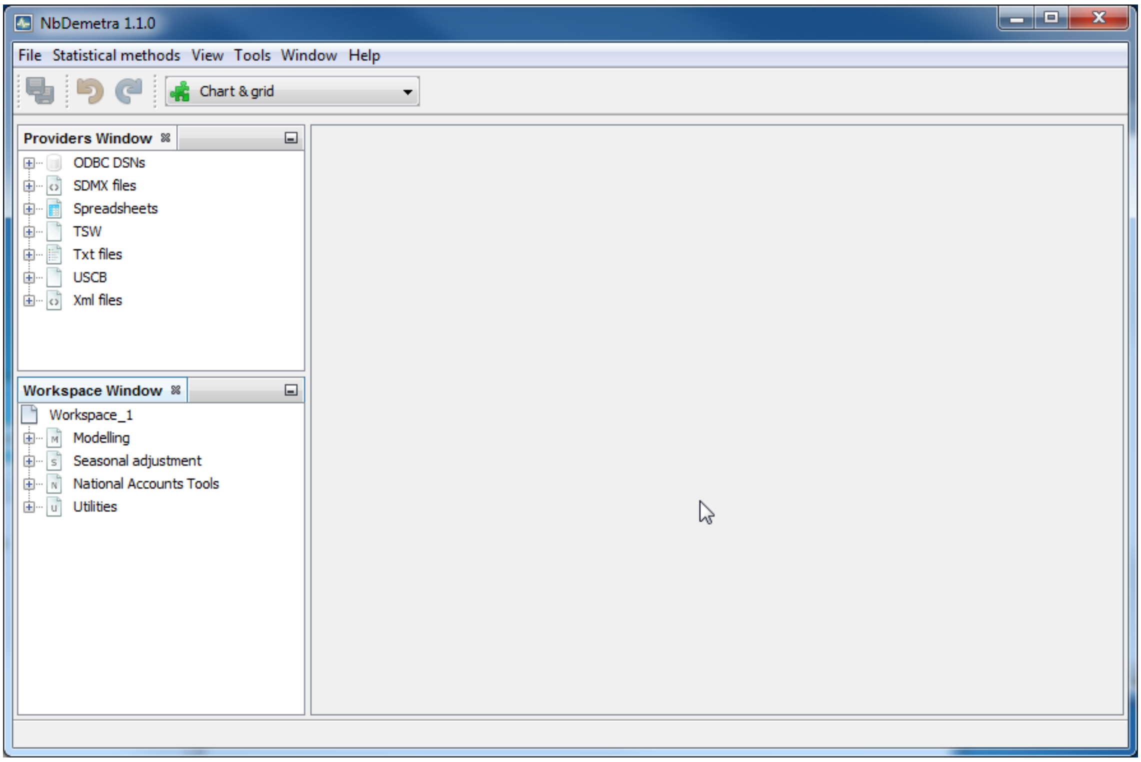


Figure 2. Batch mode.

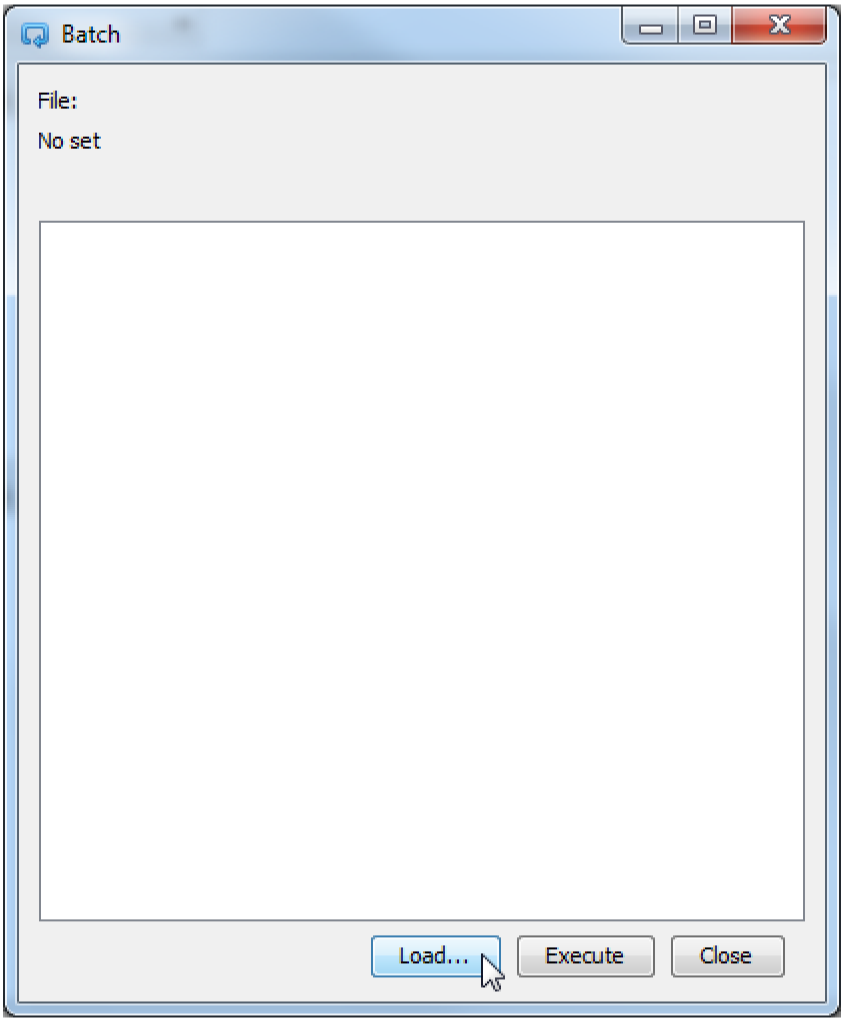


Figure 3. The Providers Window.

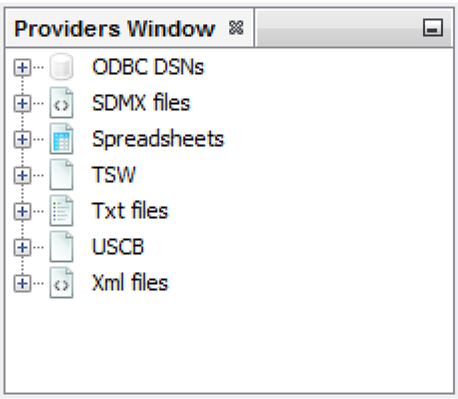


Figure 4. Workspace Window.

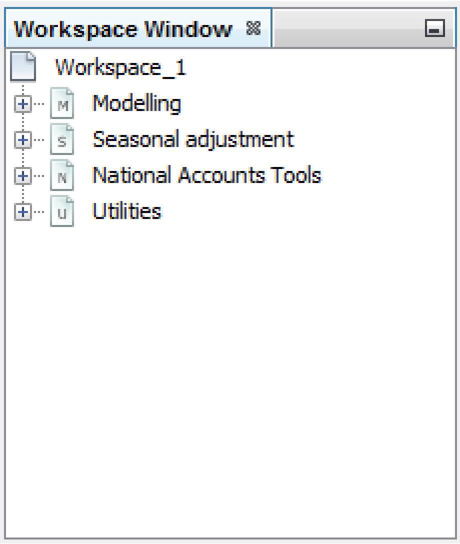
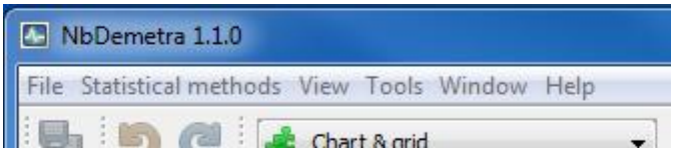


Figure 5. Main tool bar.



The "Statistical methods" pull-down menu presents several groups of methods. Only the group "National Accounts Tools" belongs to the JEcotrim package. The others are either part of JDemetra+ or have been added as other plug-ins to JDemetra+.

The JEcotrim methods can be gathered into four different families, namely: temporal disaggregation techniques, benchmarking techniques, reconciliation techniques and matrix balancing techniques.

These families of methods appear in the Workspace window, under the branch 'National Accounts Tools'.

Figure 6. Methods performed by JEcotrim.

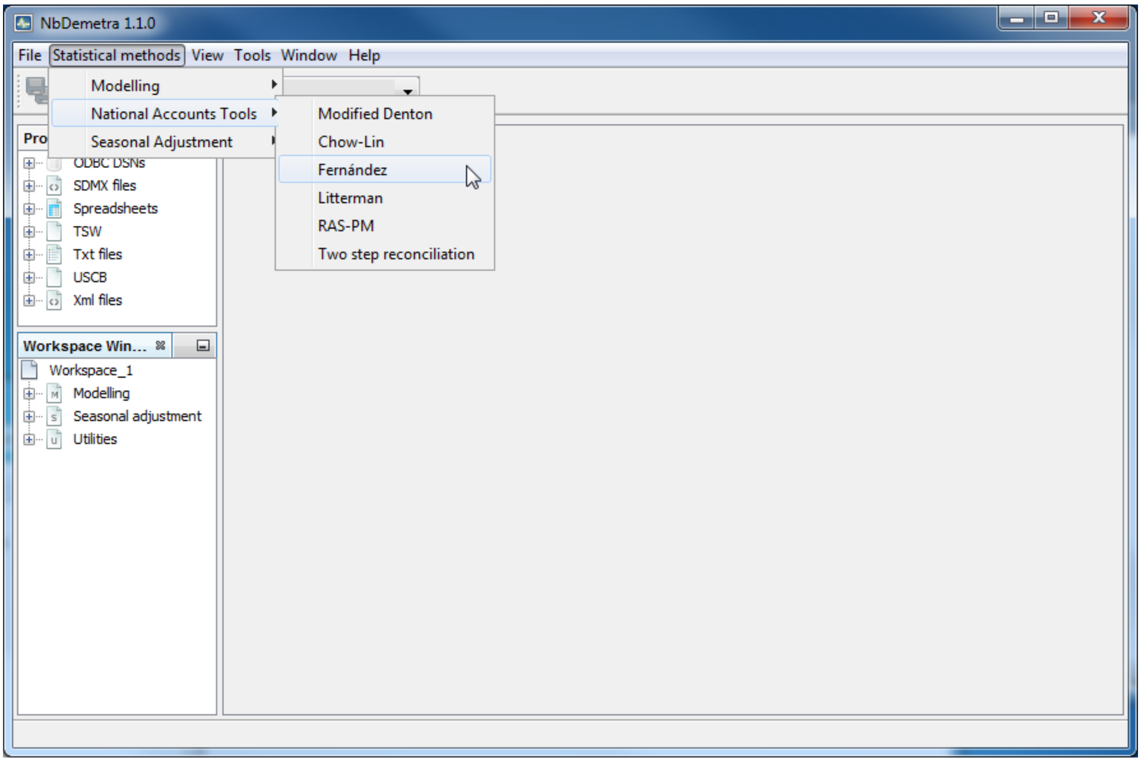
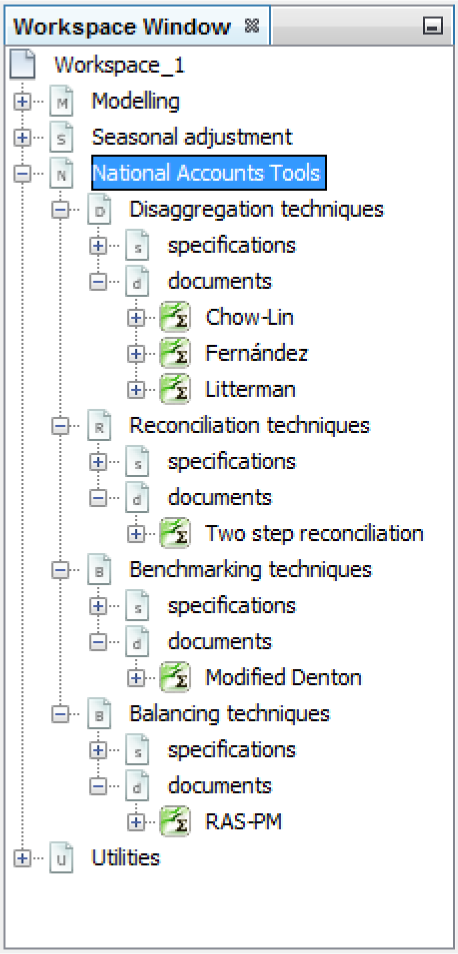


Figure 7. Families of tools


Workspace Window
Workspace_1
Modelling
Seasonal adjustment
National Accounts Tools
Disaggregation techniques
specifications
documents
Chow-Lin
Fernández
Litterman
Reconciliation techniques
specifications
documents
Two step reconciliation
Benchmarking techniques
specifications
documents
Modified Denton
Balancing techniques
specifications
documents
RAS-PM
Utilities

2.1.2. INSTALLING JECOTRIM

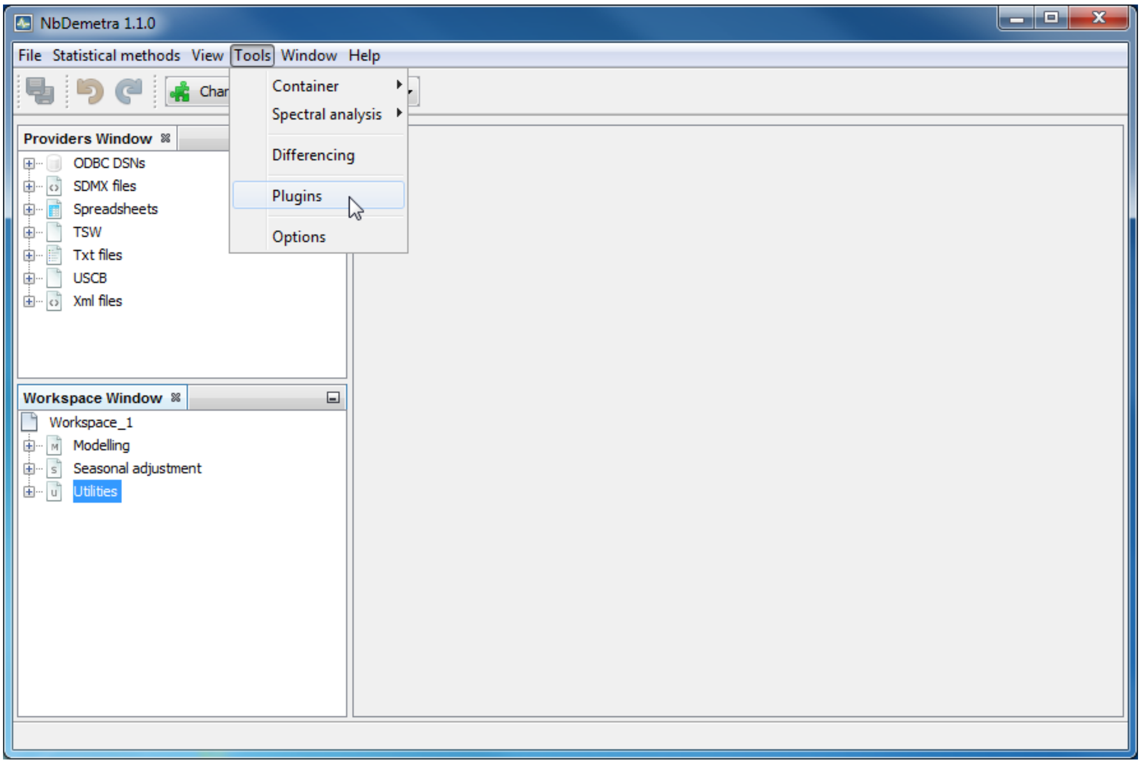
As a pre-requisite, JEcotrim needs JDemetra+ (a.k.a. NbDemetra or JDemetra or Demetra+) to be installed in the user's environment. The required version of JDemetra+ is 1.1.0.

If JDemetra+ is not installed (and working properly) or if it is not in the required version, please install the required version first. JDemetra+ can be found here: http://www.cros-portal.eu/content/jdemetra .

Once it is checked that JDemetra+ is installed in a suitable version, the installation of the JEcotrim module is done as follows.

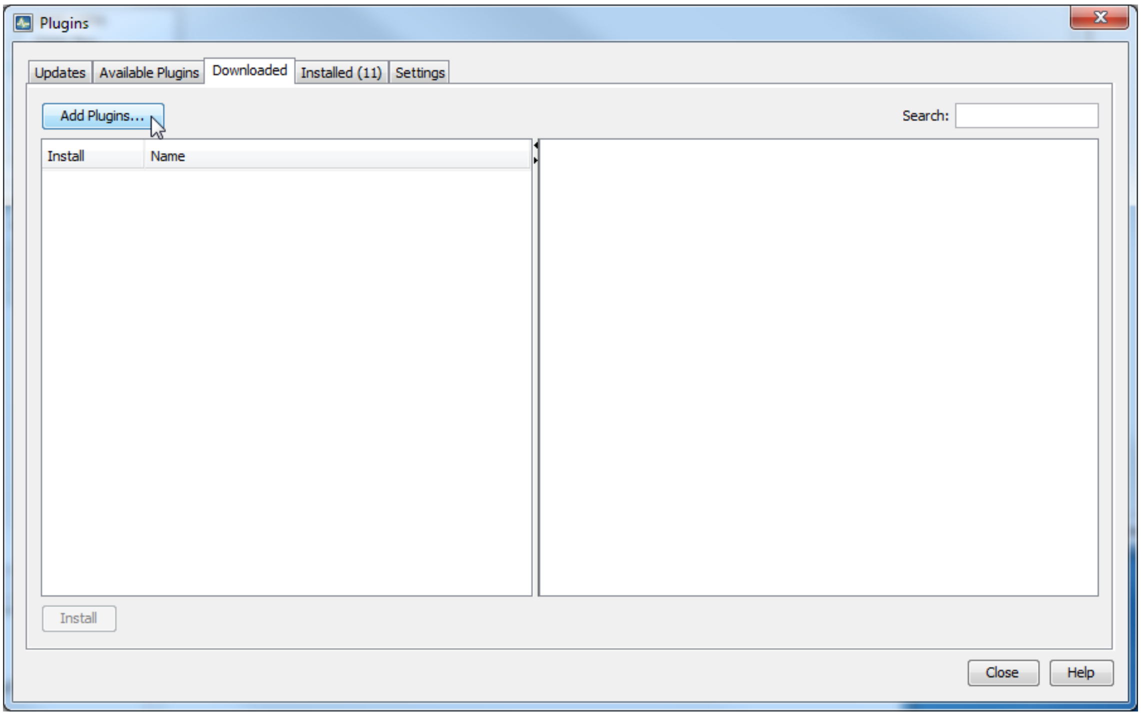
In JDemetra+, select the 'Plugins' in the menu 'Tools'.

Figure 8. Menu Plugins.



In the window 'Plugins' which is now displayed, switch to the tab 'Downloaded' and click the button 'Add Plugins…'.

Figure 9. Plugins window.

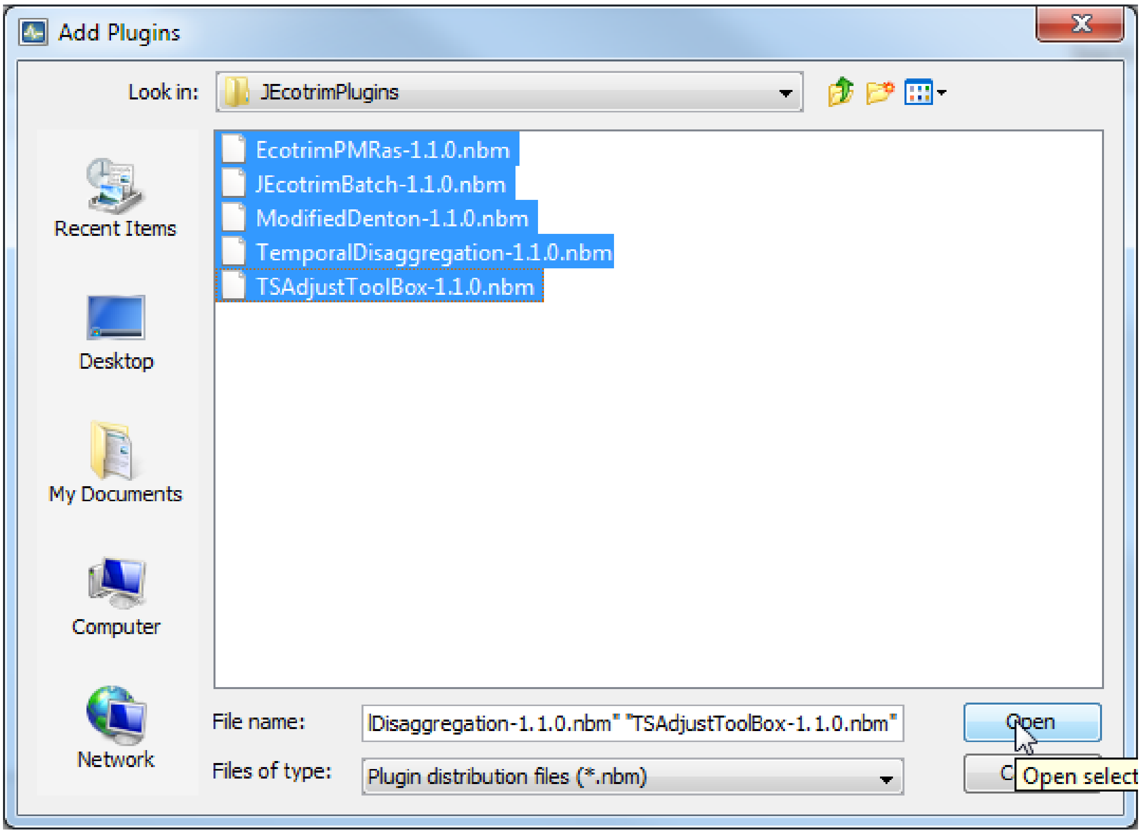


In the window 'Add Plugins' which is now displayed, browse to the directory where your JEcotrim plugins are saved, select them all then click the button 'Open'.

The plugins you have to select are the following files:

- EcotrimPMRas-1.1.0.nbm
- JEcotrimBatch-1.1.0.nbm
- ModifiedDenton-1.1.0.nbm
- TemporalDisaggregation-1.1.0.nbm
- TSAdjustToolBox-1.1.0.nbm

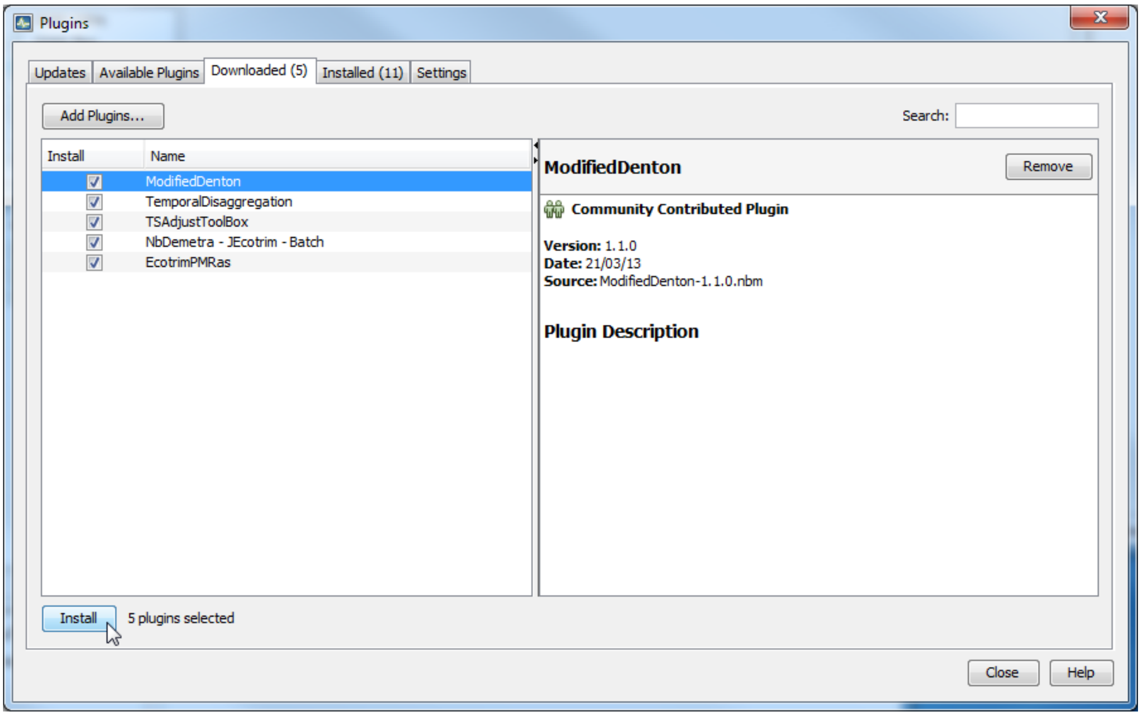
Figure 10. Add Plugins window.



The window 'Plugins' should now look as follows. In this window, click the button 'Install'.

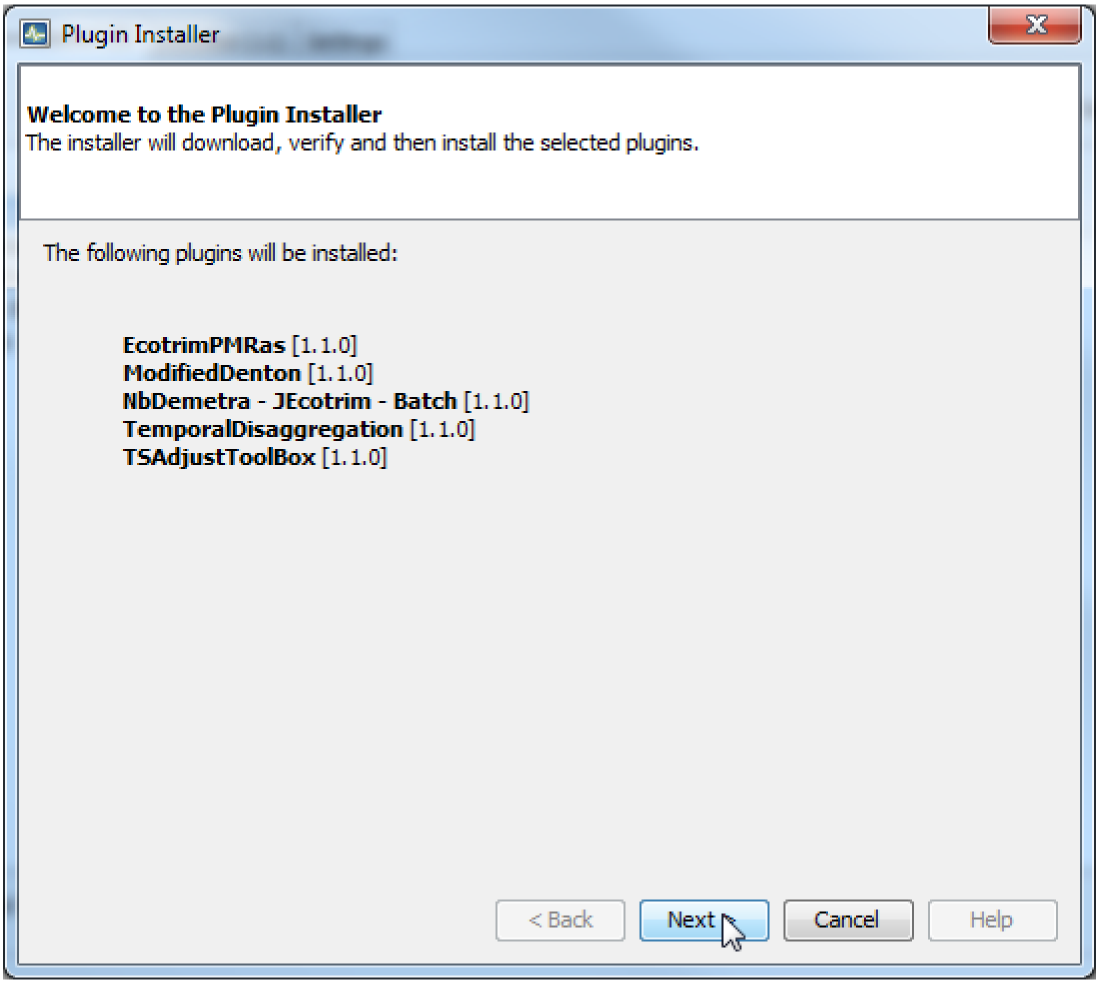
Note: in the following windows, the order of display of the plugins in the list may differ from the pictures, depending on the way you selected the plugins in the previous window. This is not a problem, since this order is indifferent.

Figure 11. Plugins window after choosing the modules.



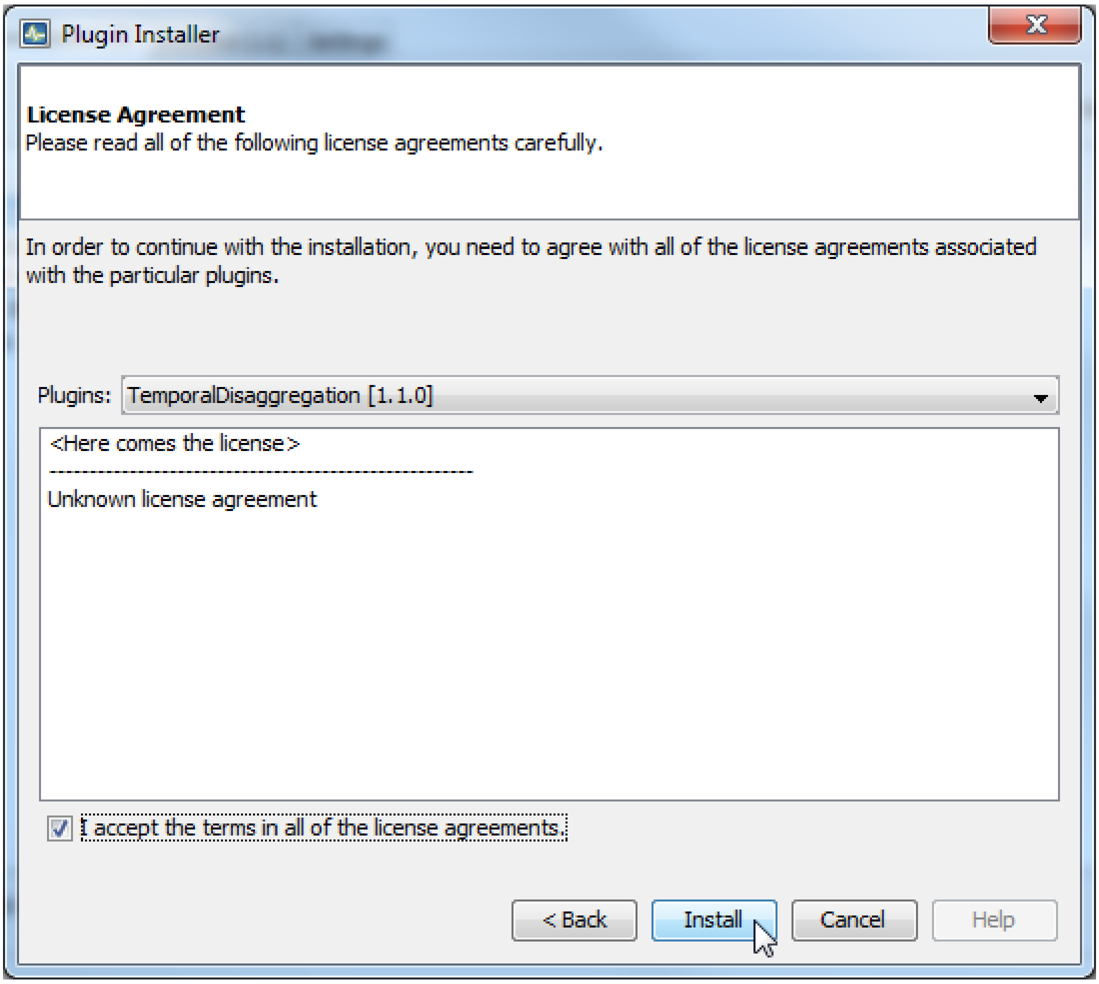
In the window 'Plugin Installer' which appears, click 'Next'.

Figure 12. Plugin installer window 1.


Plugin Installer
Welcome to the Plugin Installer
The installer will download, verify and then install the selected plugins.
The following plugins will be installed:
EcotrimPMRas [1.1.0]
ModifiedDenton [1.1.0]
NbDemetra - JEcotrim - Batch [1.1.0]
TemporalDisaggregation [1.1.0]
TSAdjustToolBox [1.1.0]
< Back
Next
Cancel
Help

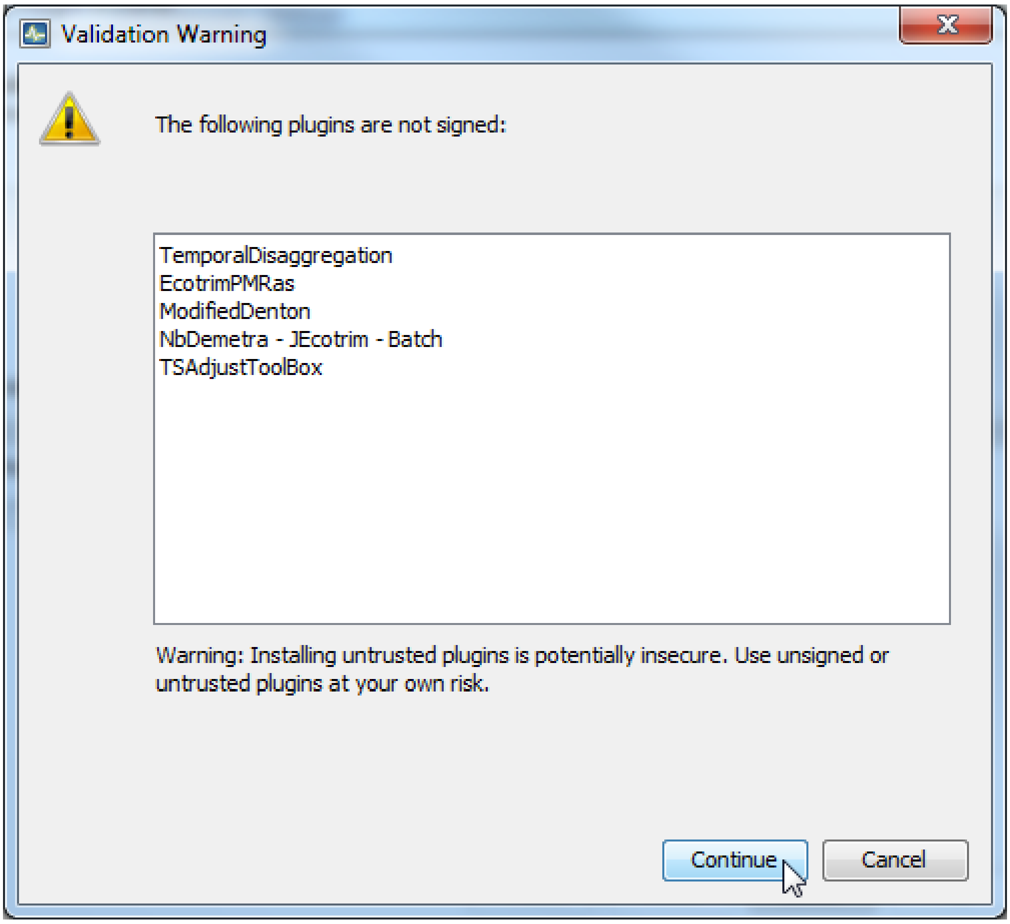
In the next window, check 'I accept the terms in all of the license agreements.' then click the button 'Install'.

Figure 13. Plugin installer window 2.



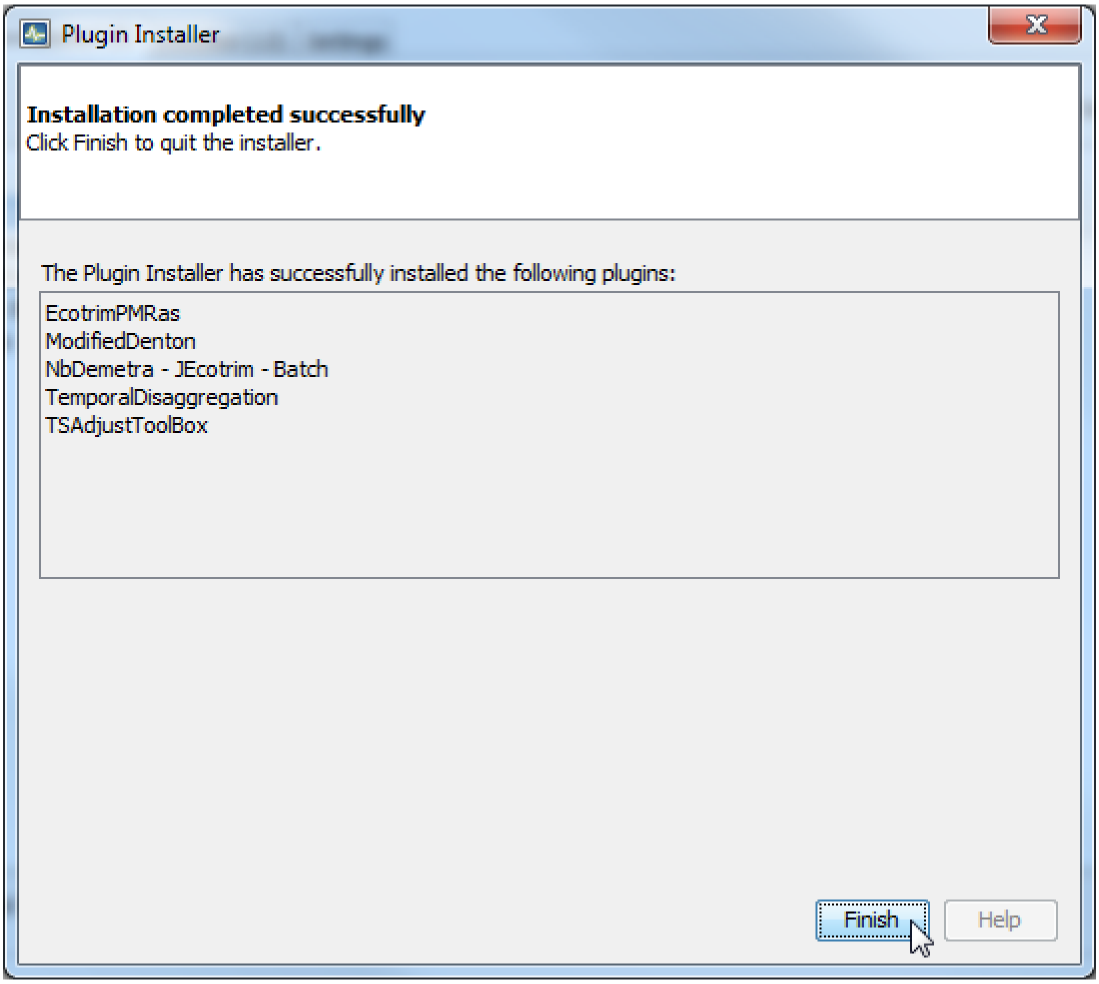
In the window 'Validation Warning', click the button 'Continue'.

Figure 14. Validation warning.



After a while, the following window should appear, telling that the installation was successful.

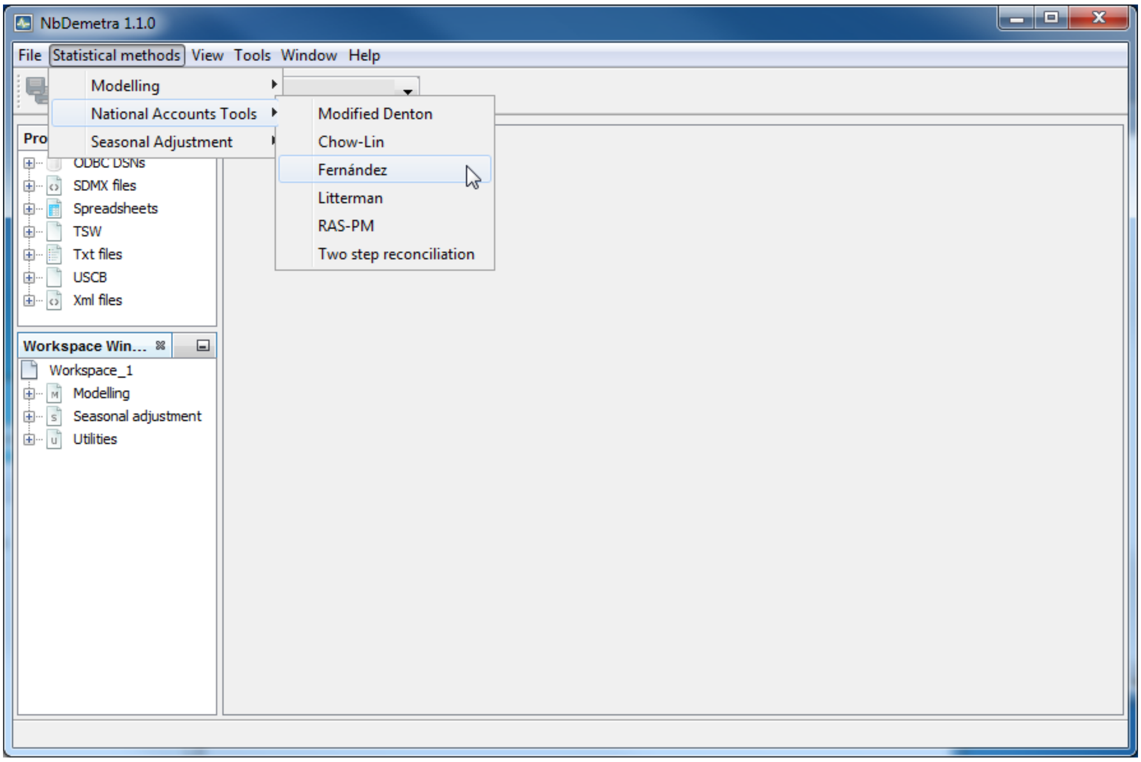
Figure 15. Plugin installer window 3



Click the button 'Finish' in this window and the button 'Close' in the window 'Plugins'.

The installation of JEcotrim module is done, as you can check in the menu 'Statistical methods'.

Figure 16. Menu Statistical methods after installation.



There is now a new menu item 'National Accounts Tools' besides the other existing modules and this menu contains 6 methods, as seen in the above picture.

Important: Restart JDemetra+ before using JEcotrim module. It will not work properly unless you restart JDemetra+.

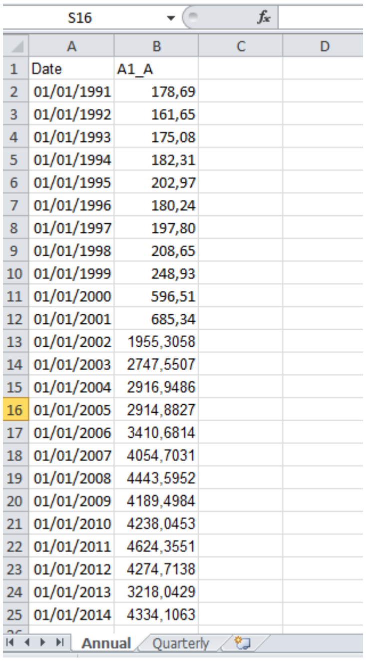
2.2. IMPORTING DATA

In this section we show how to access the data sets and of how to bring one of those into JEcotrim, because the structure of the dataset is important. It is noteworthy to say that the data input supplies the information needed for the series to be disaggregated, benchmarked, reconciliated or balanced.

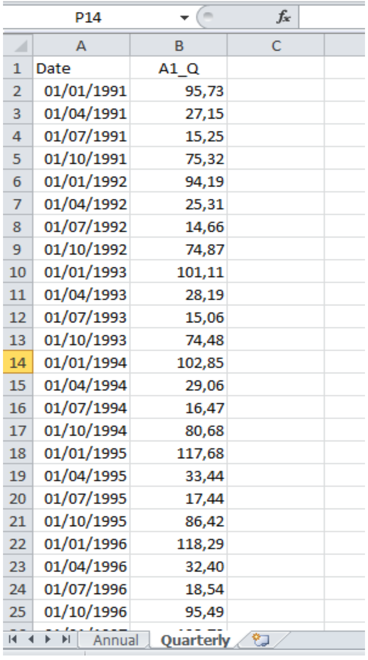
First, we will present the following Excel file called "A1.xlsx" which contains in sheet "Annual" the low frequency series (A1_A), and in sheet "Quarterly" the high frequency series (A1_Q). Both series are represented in Figure 17 and Figure 18 respectively.

ssssssssssssssssssssssss

Figure 17. A1_A series aggregated series)



	A	B
1	Date	A1_A
2	01/01/1991	178,69
3	01/01/1992	161,65
4	01/01/1993	175,08
5	01/01/1994	182,31
6	01/01/1995	202,97
7	01/01/1996	180,24
8	01/01/1997	197,80
9	01/01/1998	208,65
10	01/01/1999	248,93
11	01/01/2000	596,51
12	01/01/2001	685,34
13	01/01/2002	1955,3058
14	01/01/2003	2747,5507
15	01/01/2004	2916,9486
16	01/01/2005	2914,8827
17	01/01/2006	3410,6814
18	01/01/2007	4054,7031
19	01/01/2008	4443,5952
20	01/01/2009	4189,4984
21	01/01/2010	4238,0453
22	01/01/2011	4624,3551
23	01/01/2012	4274,7138
24	01/01/2013	3218,0429
25	01/01/2014	4334,1063



	A	B
1	Date	A1_Q
2	01/01/1991	95,73
3	01/04/1991	27,15
4	01/07/1991	15,25
5	01/10/1991	75,32
6	01/01/1992	94,19
7	01/04/1992	25,31
8	01/07/1992	14,66
9	01/10/1992	74,87
10	01/01/1993	101,11
11	01/04/1993	28,19
12	01/07/1993	15,06
13	01/10/1993	74,48
14	01/01/1994	102,85
15	01/04/1994	29,06
16	01/07/1994	16,47
17	01/10/1994	80,68
18	01/01/1995	117,68
19	01/04/1995	33,44
20	01/07/1995	17,44
21	01/10/1995	86,42
22	01/01/1996	118,29
23	01/04/1996	32,40
24	01/07/1996	18,54
25	01/10/1996	95,49

Figure 18. A1_Q series (related series).

Once the user has the series object of analysis as indicated in both figures, then process of data loading is indicated throughout Figure 19 to Figure 23. Note that the "Spreadsheets" provider is part of the JDemetra+ environment so the XLSX files must respect the rules described in the documentation of JDemetra+. Here is an overview of these rules, please refer to JDemetra+ documentation for a complete description:

– The time series can be presented in rows or in columns
– Dates within must be in the first column (A) if the time series are in columns or in the first row (1) is the times series are in rows

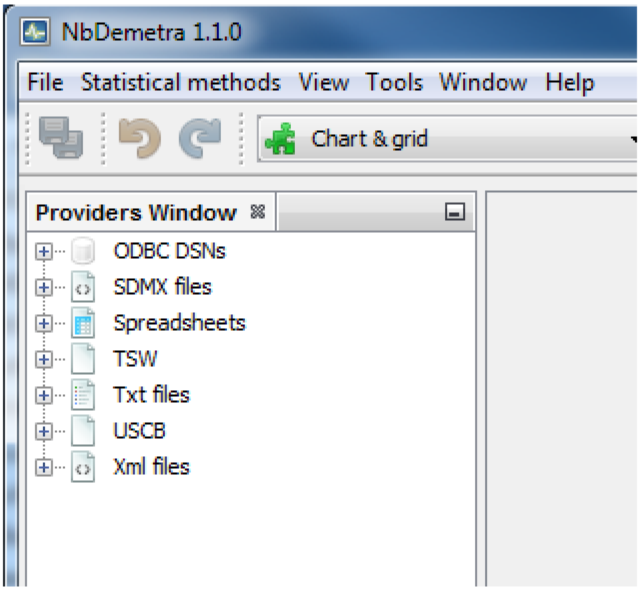
As the user will see throughout this manual, different types of input data are allowed:

– Aggregated series
– Preliminary series (related series)
– Contemporaneous constraints
– Other series

JEcotrim recognizes the frequency of the aggregated series (see Annex 1). Thus, the loading data process comprises the following steps:

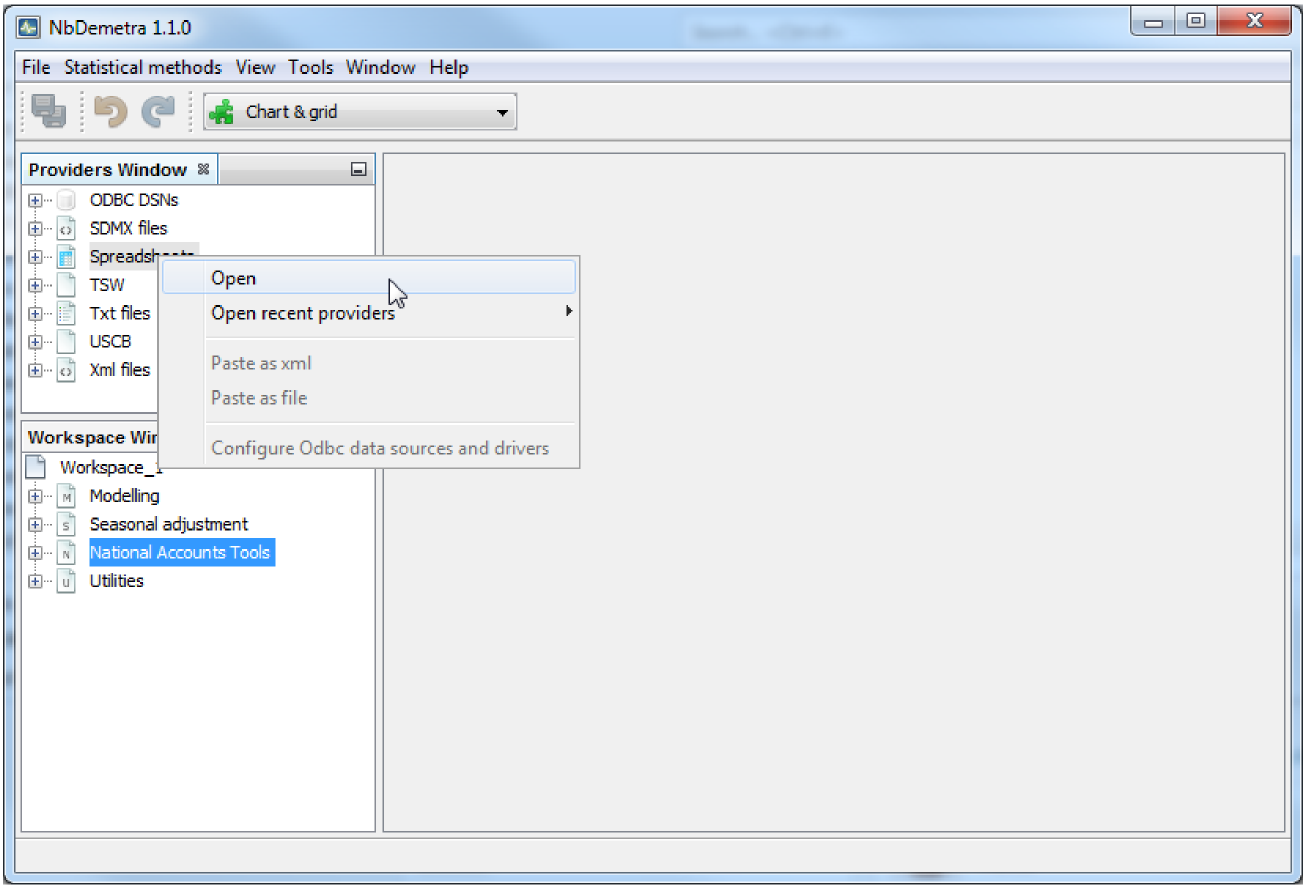
First, find the "Providers Window" in the left-hand side of JDemetra+,

Figure 19. Providers window in JDemetra+.



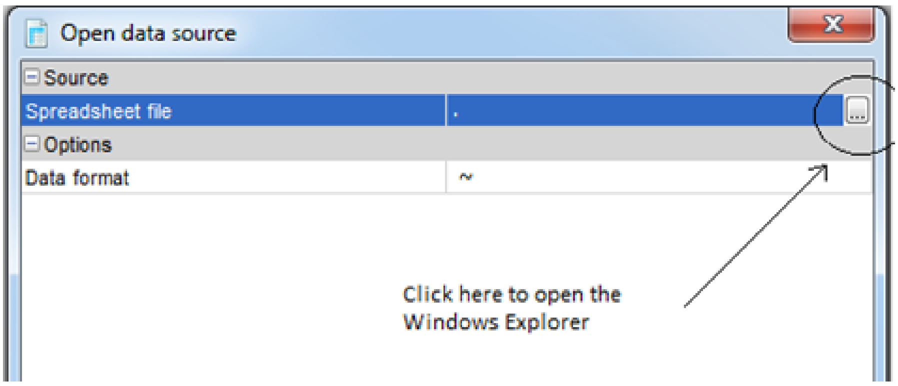
Then right-click in "Spreadsheets" and left-click (hereafter click) Open

Figure 20. First step in loading data process.



Second, in the following box, click on the button that looks like a three dots inside a box, as shown in Figure 21. If you let the "Data format" option as it is by default (value "~"), JDemetra+ "Spreadsheets" provider will expect the dates to be in the date format of MS Excel. Please refer to JDemetra+ documentation to learn more about the "Data format" option.

Figure 21. Second step in loading dataset.



Find the data object of analysis in the hard drive disk and click OK. Once the series has been selected, it appears in the box as shown in Figure 22. Finally, click on the OK button to load the end process of loading series, in the present case, A1. At this point, the JEcotrim should display the series (A1) as indicated in Figure 23.

Figure 22: Second to last step of loading dataset.

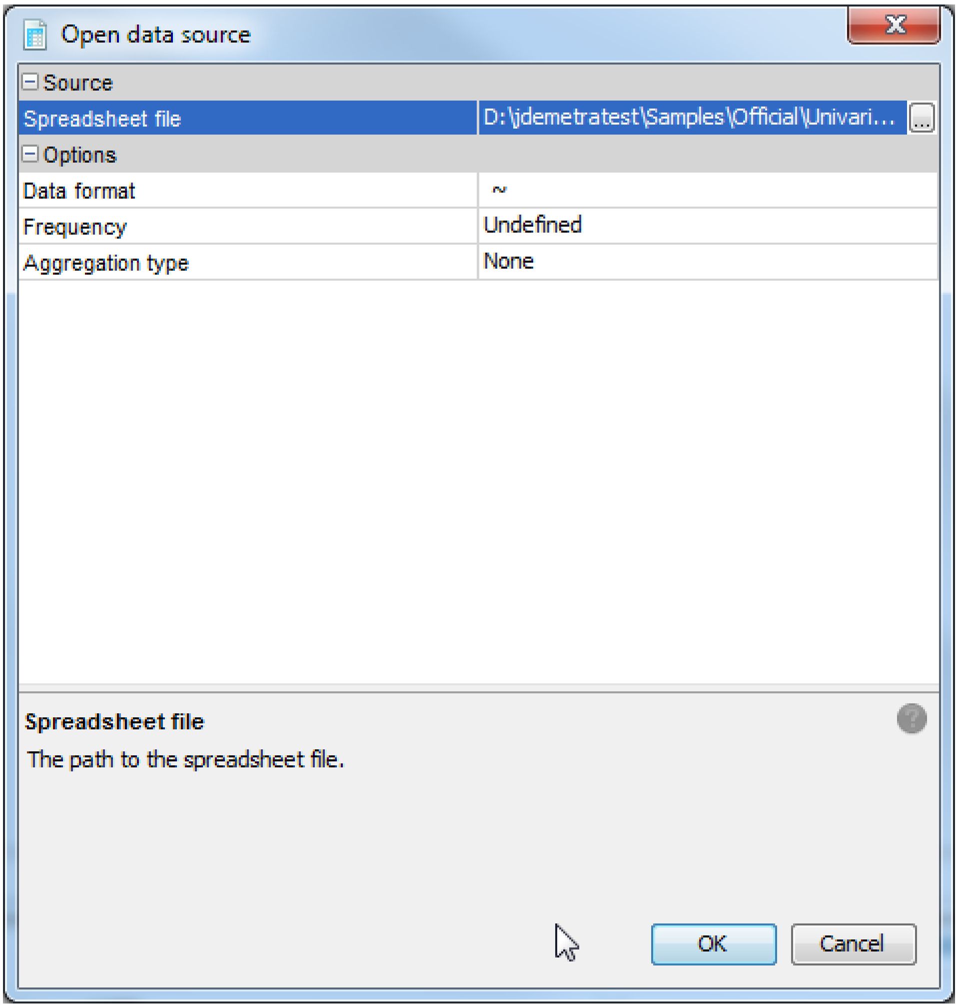
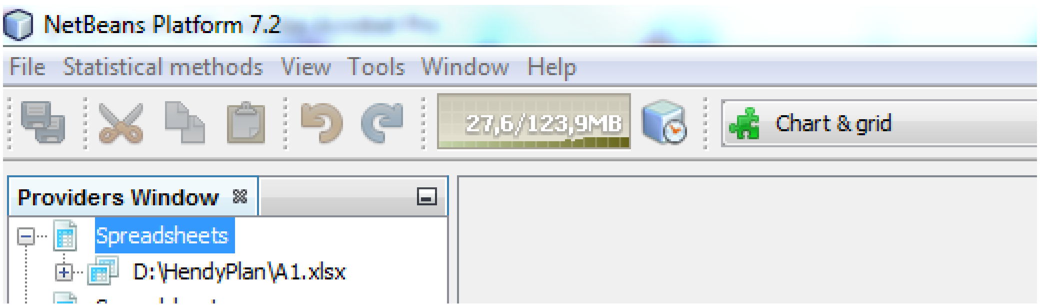


Figure 23. Final view of loaded dataset.



2.3. SESSION IN JECOTRIM

A typical session in JEcotrim involves:

- an initial data acquisition stage, that is, the data the user is employing as input. At this point, the user must recall that the data can be loaded before or after the statistical method is selected, the order does not matter.
- Next, the estimation stage of the available information depending upon the method selected and the specifications given by the user.
- A final stage in which the outputs are presented in the same Workspace.

Note: JEcotrim gives the option to the user to copy the results and export to any data processor.

Before leaving the program asks confirmation of saving the current Workspace. After that, the user leaves the program.

3. JECOTRIM INTERACTIVE MODE

In this chapter, the user is introduced in a first stage, to the handling of the univariate temporal disaggregation methods, namely, Chow-Lin, Fernández and Litterman; in a second stage, to the univariate benchmarking technique (Modified Denton); next, to a reconciliation method (the Two-step procedure) and finally, to a matrix balancing procedure (RAS-PM).

3.1. TIME SERIES DISAGGREGATION TECHNIQUES

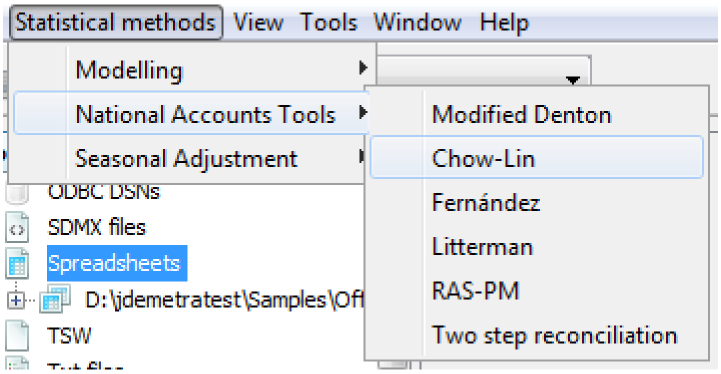
Temporal disaggregation is a process used to disaggregate a given economic time series (usually known as total, aggregated or low-frequency data) to produce values for shorter sub periods (known as disaggregated or high-frequency data). This kind of problem is often faced by national statistical institutes, for example when producing Quarterly National Accounts series by means of indirect techniques.

In the last decades, these techniques have been widely investigated in the literature. Thus, two main different approaches can be distinguished: first, methods that do not make use of the information from related series but rely in time series models to derive a smooth path for the unobserved series; and second, methods that do use related series. The methodology for the first approach can be found in Stram and Wei (1986) and Wei and Stram (1990), whilst the second group of techniques is described in Chow and Lin (1971), Fernández (1981) and Litterman (1983), among others. This guide is focused on the latter group.

3.1.1. CHOW AND LIN (1971)

This temporal disaggregation technique, based on regression, works out a least-squares optimal solution on the basis of a linear regression model involving the quarterly aggregate series and the related quarterly series and having assumed an AR(1) error structure in that regression model. Thus, the user should proceed as follows: first, select Statistical methods>National Accounts Tools>Chow-Lin from the pull-down menu, as the user can see below

Figure 24. Chow-Lin method selection.



Once the method is selected, the interface shows Figure 25 with three separate parts. In the left-hand side of the screen of the interface the user can observe a drop down menu where now it appears “Summary”, “Input” and further, will find the “Output” menu as well. In the center, one can find the window which brings a complete description of the method and further, will show data, grids and charts, while in the right-hand side, the specifications menu can be found. We will explain this point further.

Figure 25. Overview of the main window for Chow-Lin technique.

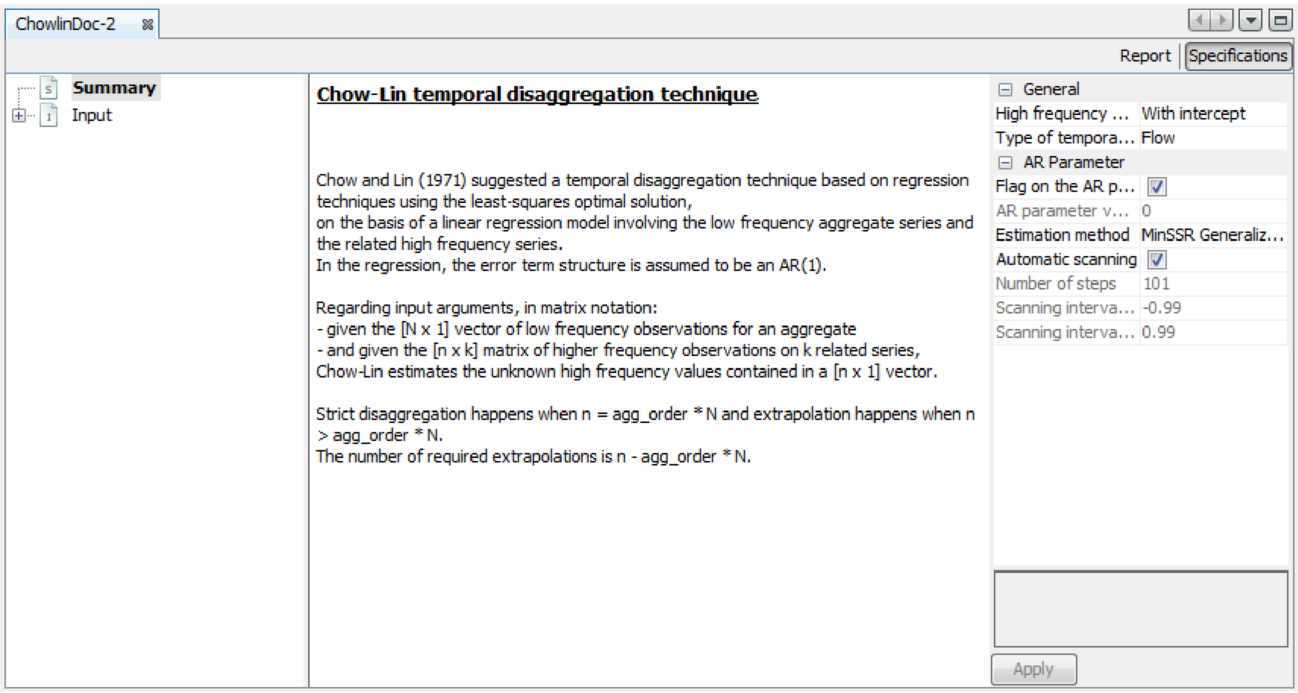


Figure 26. Detailed view of the left-hand side of Chow-Lin's window.

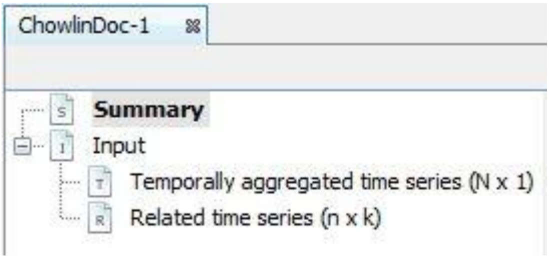
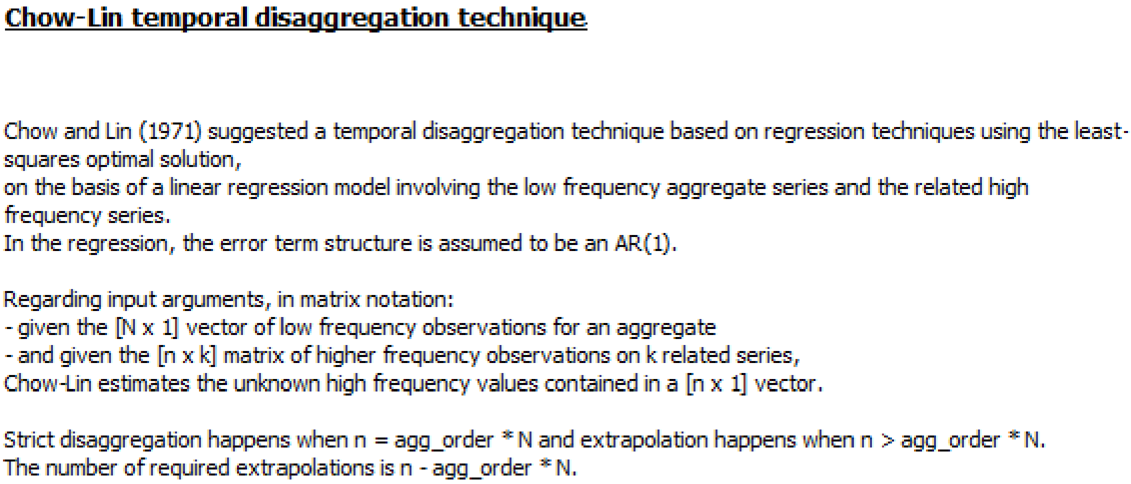


Figure 27. Detailed view of the central window of Chow-Lin's.



Chow-Lin temporal disaggregation technique

Chow and Lin (1971) suggested a temporal disaggregation technique based on regression techniques using the least-squares optimal solution,
on the basis of a linear regression model involving the low frequency aggregate series and the related high frequency series.
In the regression, the error term structure is assumed to be an AR(1).

Regarding input arguments, in matrix notation:
- given the [N x 1] vector of low frequency observations for an aggregate
- and given the [n x k] matrix of higher frequency observations on k related series,
Chow-Lin estimates the unknown high frequency values contained in a [n x 1] vector.

Strict disaggregation happens when n = agg_order * N and extrapolation happens when n > agg_order * N.
The number of required extrapolations is n - agg_order * N.

In order to display the Specifications, as shown in the picture below, you have to click on the button 'Specifications' in the upper-right part of the window.

Figure 28. Overview of the "Specifications" menu for Chow-Lin's.

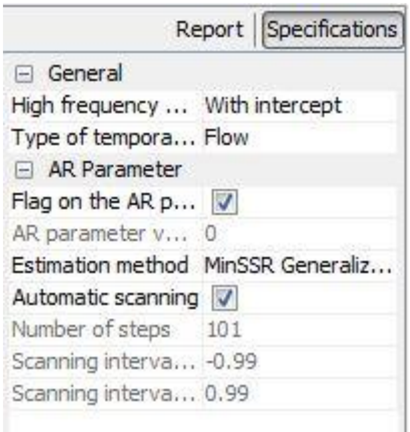
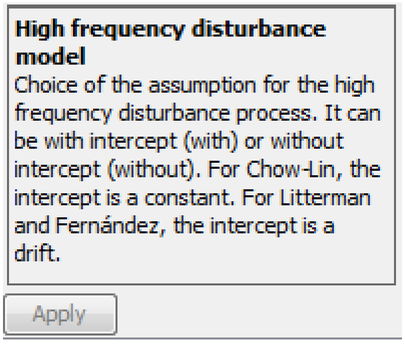


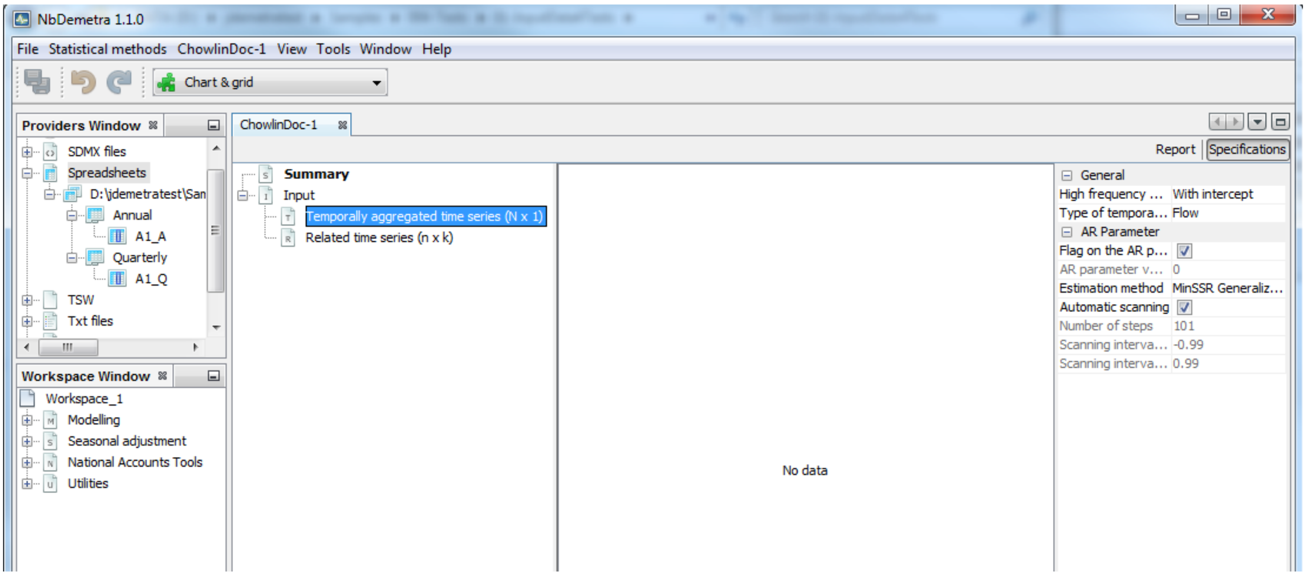
Figure 29. Detailed view of the "Specifications" menu which contains its legend.



After presenting the workspace for Chow-Lin method, we proceed to explain the data loading process[2]:

First, select Input>Temporal aggregated time series and drag the "Annual" series and drop in the main workspace window, where there is no data yet, as indicated in Figure 30.

Figure 30. Loading aggregated series

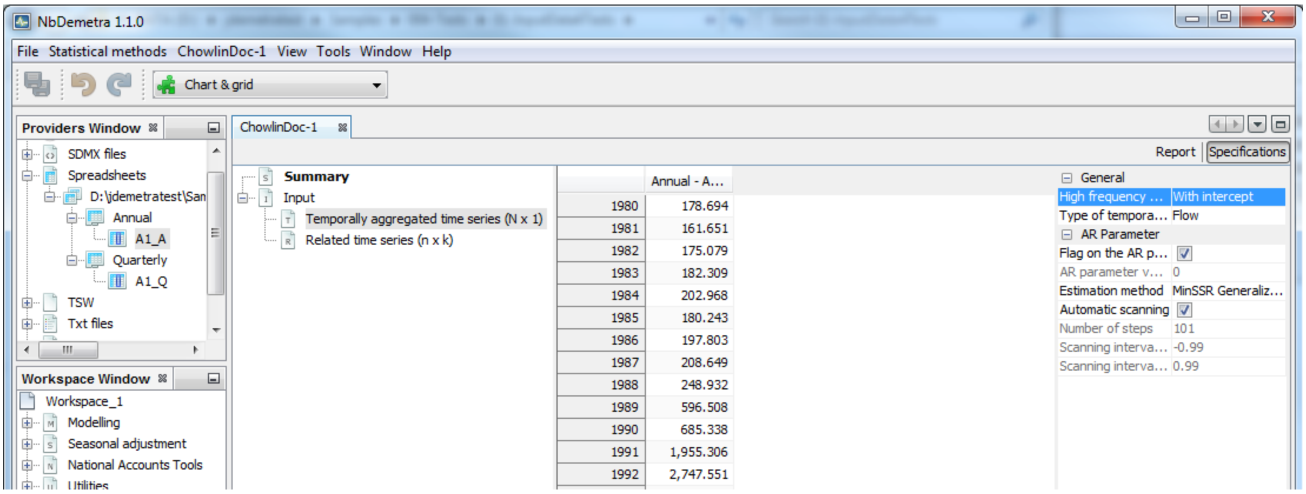


If the task has been done correctly, then the data should appear in the main workspace window as it can be seen next in Figure 31.

---

[2] As mentioned previously, the data loading can be done after the statistical method has been selected. In this guide we present the cases selecting the data after having chosen the method.

Figure 31. Annual series to be disaggregated.



The same procedure has to be followed so as to load the related series (see Figure 32 and Figure 33).

Figure 32. Loading related series.

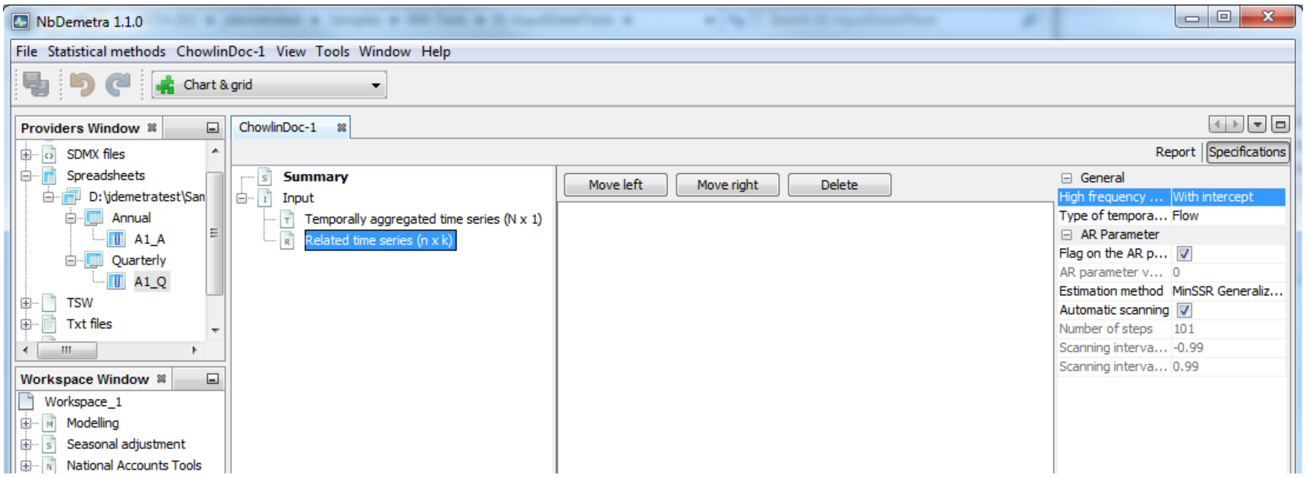
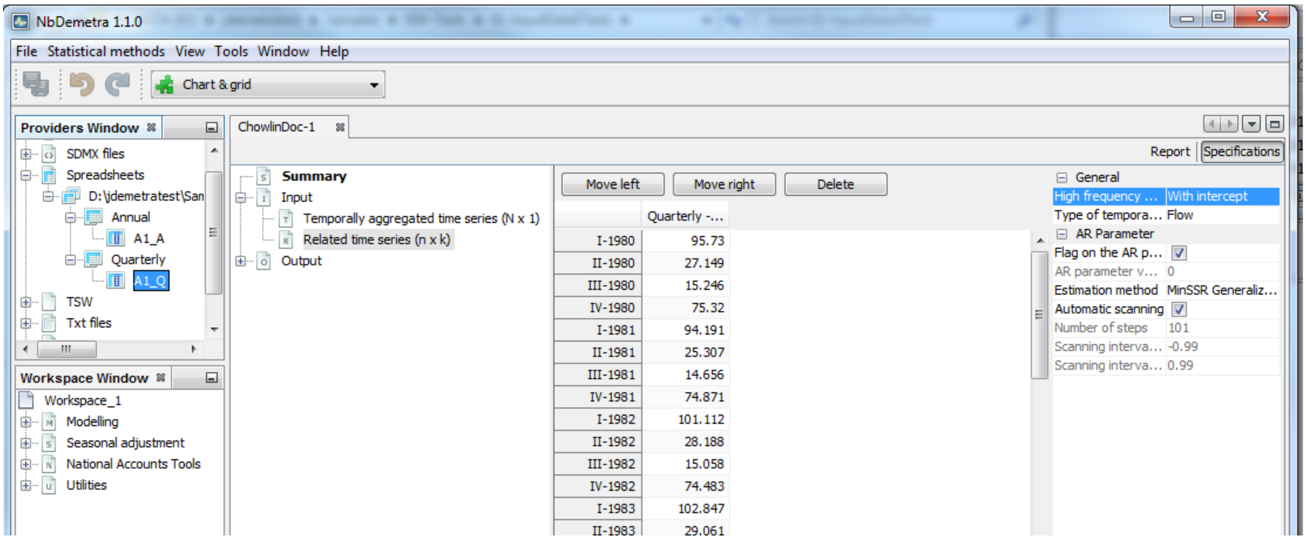
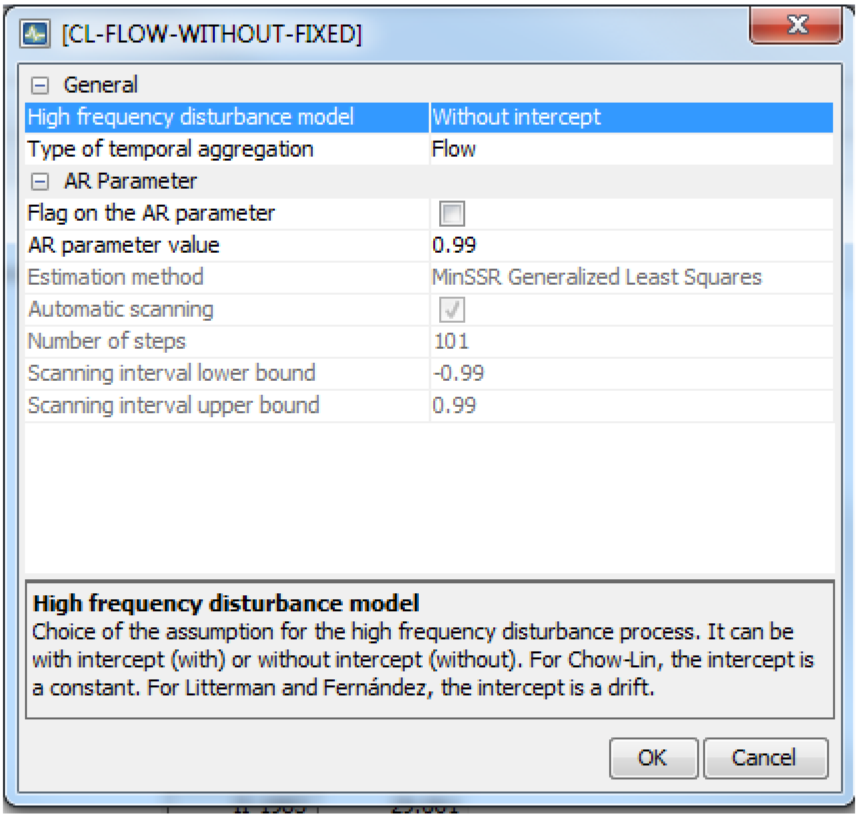


Figure 33. Related series loaded.



At this point, it is time to enter the "Specifications" menu. This menu is shown just below

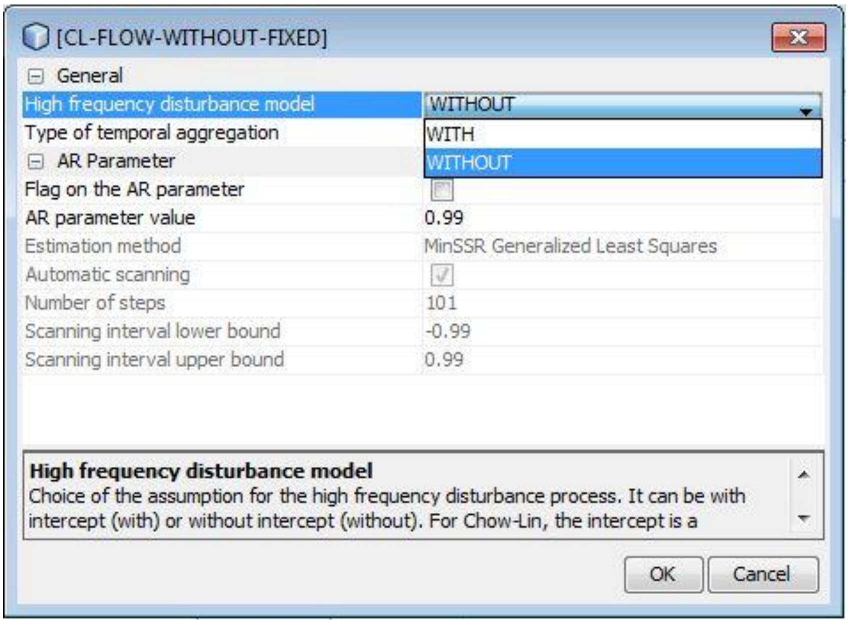
Figure 34. Specifications menu.



The interface gives the option to the user to:

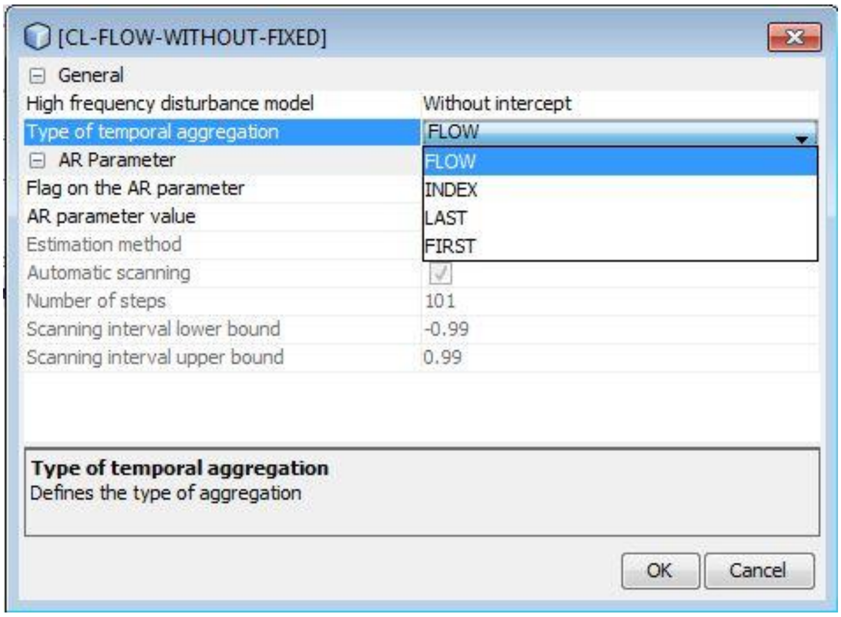
– Choose whether or not to include an intercept term in the regression model.

Figure 35. Constant term included or not in the model.



– Choose among four types of temporal aggregation: flow, index, first and last (see Annex 2).

Figure 36. Selection of the type of the temporal aggregation.



– Leave the autoregressive parameter fixed at user's choice (-.999 till 0.999) or leave it to be estimated according to two different methods.

Figure 37. Decision about AR parameter (estimated).

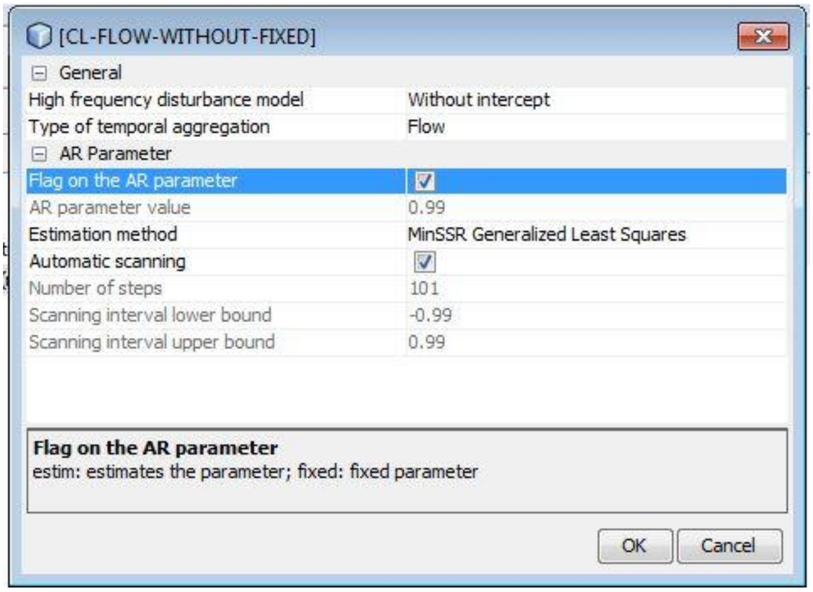
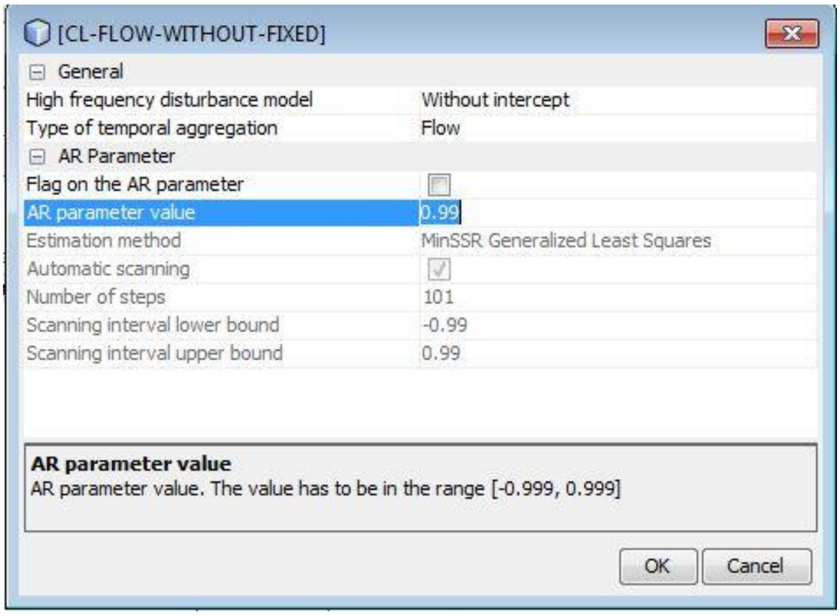
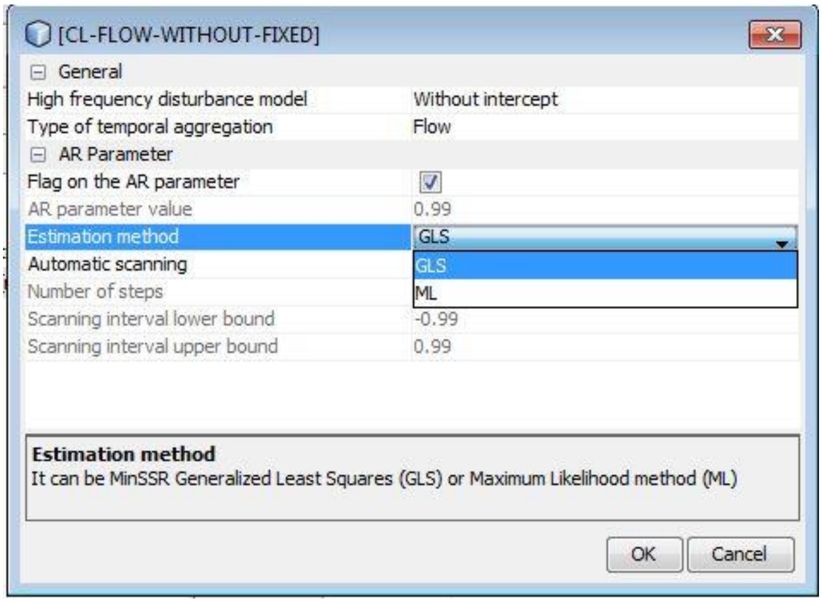


Figure 38. Decision about AR parameter (fixed by user).



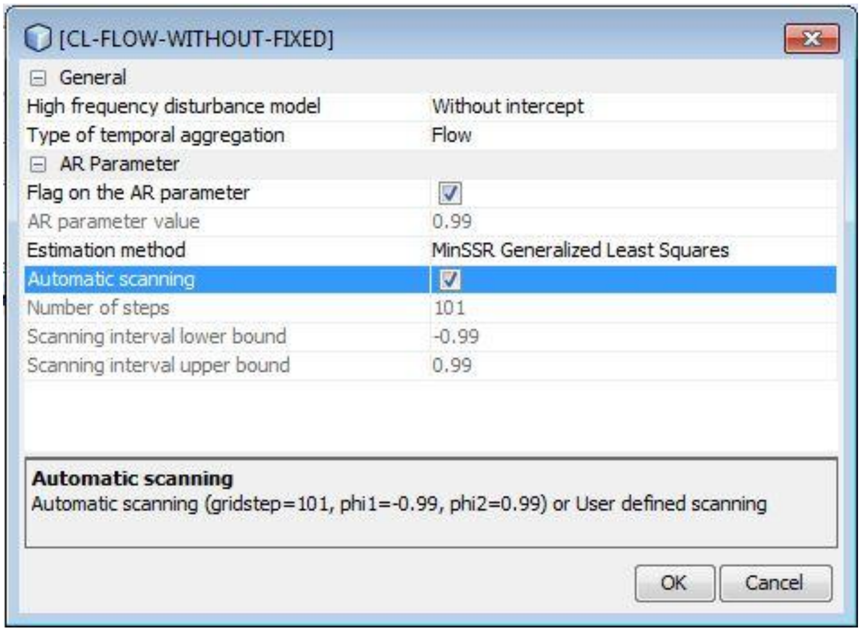
– Select the estimation method: min SSR GLS or maximum likelihood (ML).

Figure 39. Decision about the estimation method.



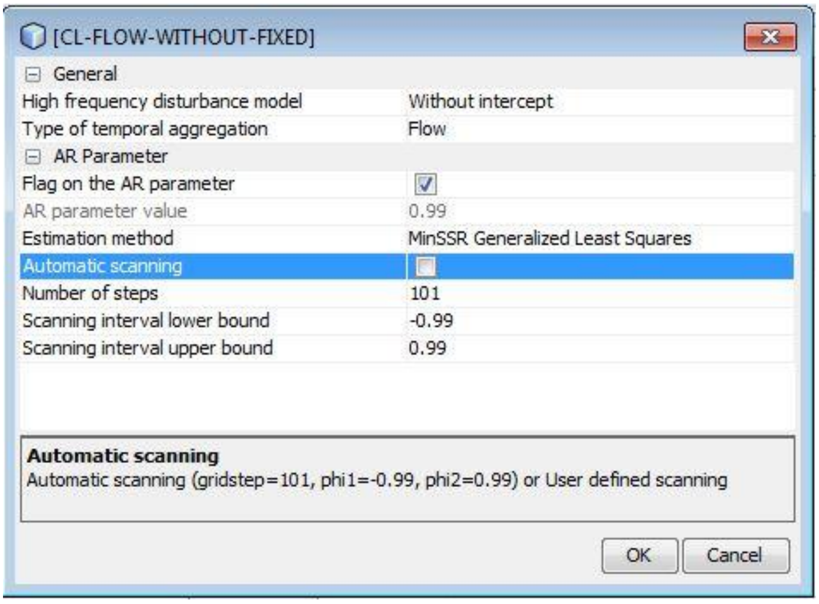
– Decide about the scanning procedure, whether leave it with pre-defined settings (101 steps scanning in an interval ranging from -0.99 to 0.99) or enter user-defined number of steps and scanning interval bounds.

Figure 40. Automatic scanning procedure.



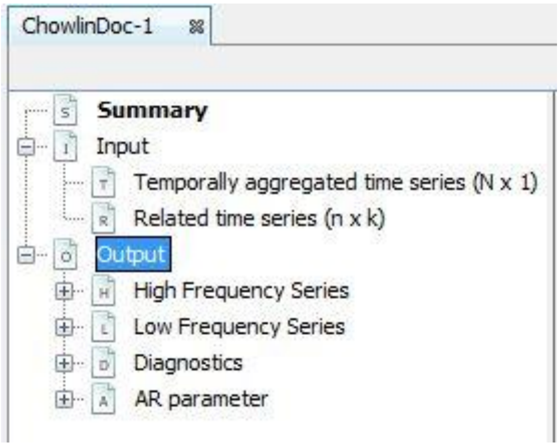
– Set the number of steps and the lower and upper bounds for the scanning procedure if it has been selected user-defined rather than automatic.

Figure 41. Fixed scanning procedure (by user).



After having made the corresponding modifications within the specifications and validating them by clicking 'OK' or 'Apply' (depending on the specifications window where you do the modifications), JEcotrim gives the results automatically in a straight-forward manner[3]:

Figure 42. Output menu (Chow-Lin's).



Then clicking in "High Frequency Series" and "Grid" provides:

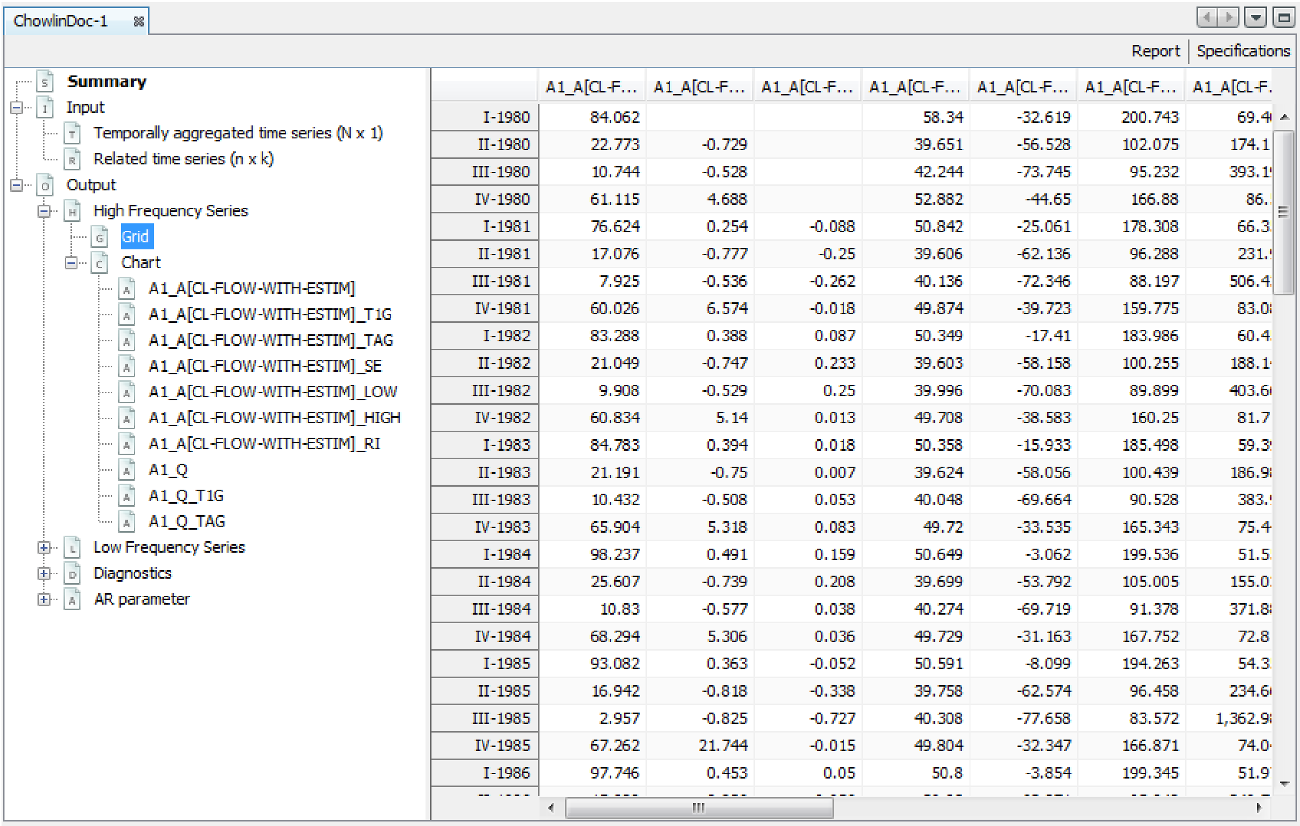
- the disaggregated estimated time series (1st column A1_A[CL-FLOW-WITH-ESTIM])
- the lower and upper limits of the confidence interval (5th and 6th columns A1_A[CL-FLOW-WITH-ESTIM]_LOW and A1_A[CL-FLOW-WITH-ESTIM]_HIGH, respectively)
- the standard error of the estimates (4th column A1_A[CL-FLOW-WITH-ESTIM]_SE)
- the reliability indicators (7th column A1_A[CL-FLOW-WITH-ESTIM]_RI)
- the lag 1 growth rate (2nd column A1_A[CL-FLOW-WITH-ESTIM]_T1G)

[3] Notice that if the user does not change the specifications, the program returns the results as if everything has been done following the guidelines proposed.

– the annual growth rate (3rd column A1_A[CL-FLOW-WITH-ESTIM]_TAG)
– the original related time series (8th column A1_Q)
– the lag 1 growth rate of the related time series (9th column A1_Q_T1G)
– the annual growth rate of the related time series (10th column A1_Q_TAG)

See Annex 3 for details about the confidence interval, the reliability indicator and the growth rates.

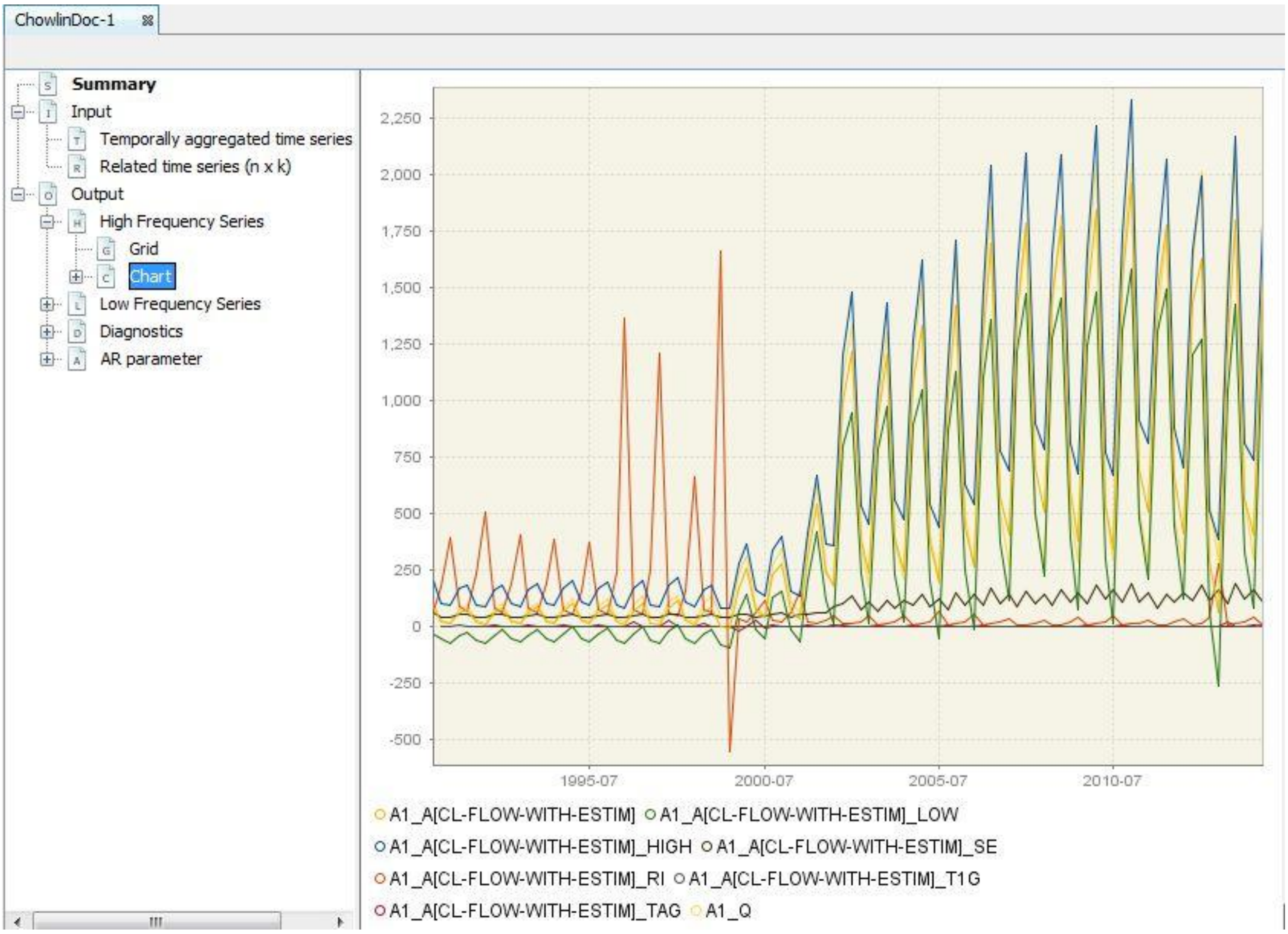
Figure 43. Grid that provides the disaggregated estimated series. (Chow-Lin's).



	A1_A[CL-F...	A1_A[CL-F...	A1_A[CL-F...	A1_A[CL-F...	A1_A[CL-F...	A1_A[CL-F...	A1_A[CL-F.
I-1980	84.062			58.34	-32.619	200.743	69.4
II-1980	22.773	-0.729		39.651	-56.528	102.075	174.1
III-1980	10.744	-0.528		42.244	-73.745	95.232	393.1
IV-1980	61.115	4.688		52.882	-44.65	166.88	86.
I-1981	76.624	0.254	-0.088	50.842	-25.061	178.308	66.3
II-1981	17.076	-0.777	-0.25	39.606	-62.136	96.288	231.
III-1981	7.925	-0.536	-0.262	40.136	-72.346	88.197	506.4
IV-1981	60.026	6.574	-0.018	49.874	-39.723	159.775	83.0
I-1982	83.288	0.388	0.087	50.349	-17.41	183.986	60.4
II-1982	21.049	-0.747	0.233	39.603	-58.158	100.255	188.1
III-1982	9.908	-0.529	0.25	39.996	-70.083	89.899	403.6
IV-1982	60.834	5.14	0.013	49.708	-38.583	160.25	81.7
I-1983	84.783	0.394	0.018	50.358	-15.933	185.498	59.3
II-1983	21.191	-0.75	0.007	39.624	-58.056	100.439	186.9
III-1983	10.432	-0.508	0.053	40.048	-69.664	90.528	383.
IV-1983	65.904	5.318	0.083	49.72	-33.535	165.343	75.4
I-1984	98.237	0.491	0.159	50.649	-3.062	199.536	51.5
II-1984	25.607	-0.739	0.208	39.699	-53.792	105.005	155.0
III-1984	10.83	-0.577	0.038	40.274	-69.719	91.378	371.8
IV-1984	68.294	5.306	0.036	49.729	-31.163	167.752	72.8
I-1985	93.082	0.363	-0.052	50.591	-8.099	194.263	54.3
II-1985	16.942	-0.818	-0.338	39.758	-62.574	96.458	234.6
III-1985	2.957	-0.825	-0.727	40.308	-77.658	83.572	1,362.9
IV-1985	67.262	21.744	-0.015	49.804	-32.347	166.871	74.0
I-1986	97.746	0.453	0.05	50.8	-3.854	199.345	51.9

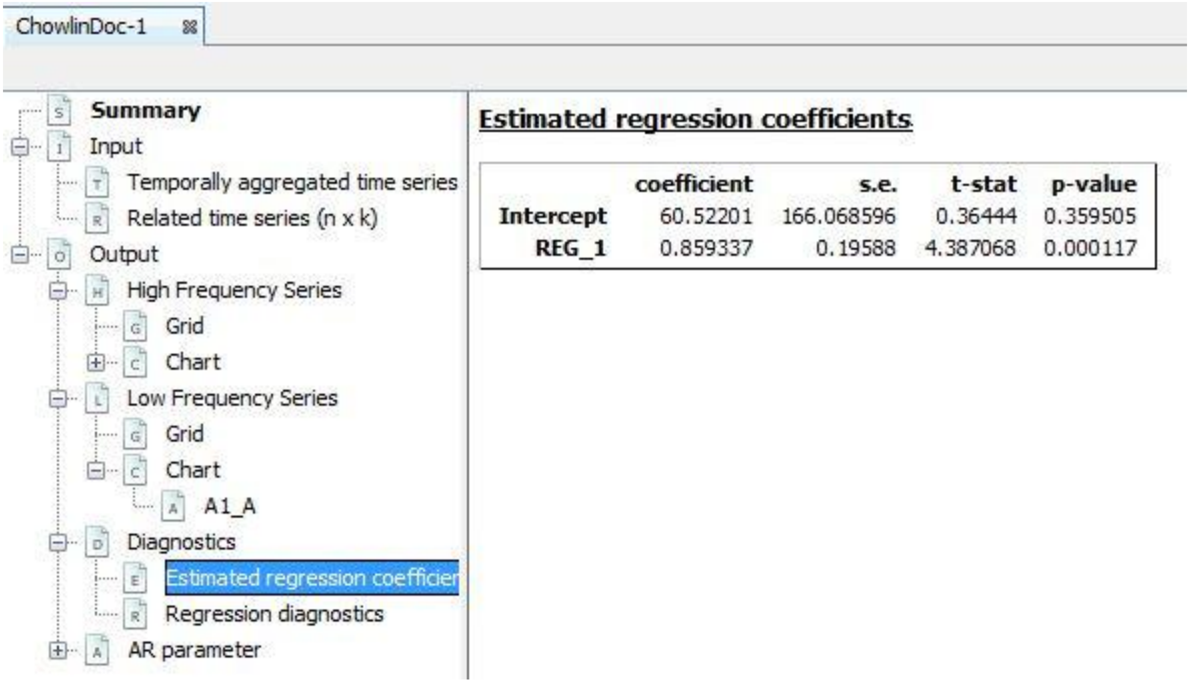
Likewise, the "Chart" offers a graphical representation of the disaggregated time series.

Figure 44. Graphical view of the previous "Grid" (Chow-Lin's).



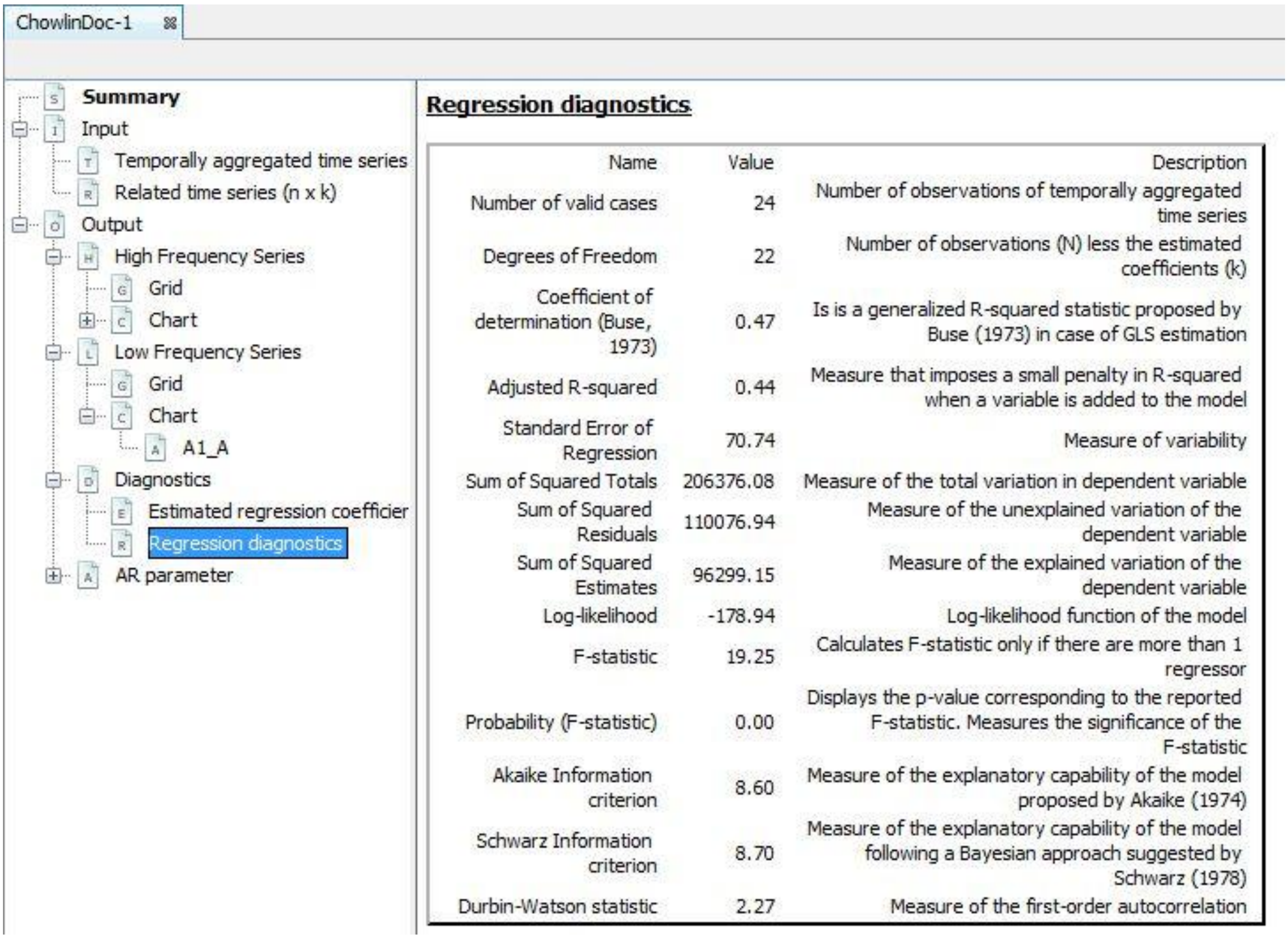
In the "Diagnostics" option, the values of the estimated parameters of the regression model are accessible.

Figure 45. Estimated regression coefficients (Chow-Lin's).



Also available is the summary of the statistics considered assessing the goodness and validity of the model (with its corresponding description to ease the work to the user).

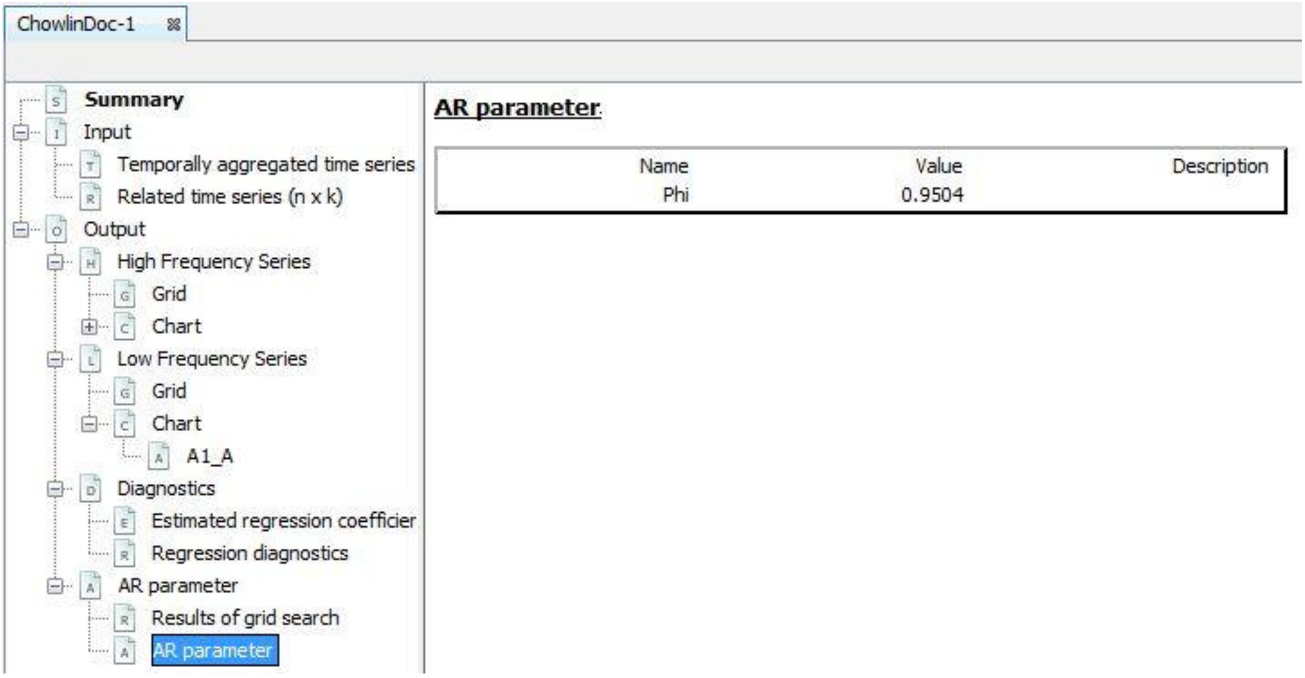
Figure 46. Summary of statistics (Chow-Lin's).



Name	Value	Description
Number of valid cases	24	Number of observations of temporally aggregated time series
Degrees of Freedom	22	Number of observations (N) less the estimated coefficients (k)
Coefficient of determination (Buse, 1973)	0.47	Is is a generalized R-squared statistic proposed by Buse (1973) in case of GLS estimation
Adjusted R-squared	0.44	Measure that imposes a small penalty in R-squared when a variable is added to the model
Standard Error of Regression	70.74	Measure of variability
Sum of Squared Totals	206376.08	Measure of the total variation in dependent variable
Sum of Squared Residuals	110076.94	Measure of the unexplained variation of the dependent variable
Sum of Squared Estimates	96299.15	Measure of the explained variation of the dependent variable
Log-likelihood	-178.94	Log-likelihood function of the model
F-statistic	19.25	Calculates F-statistic only if there are more than 1 regressor
Probability (F-statistic)	0.00	Displays the p-value corresponding to the reported F-statistic. Measures the significance of the F-statistic
Akaike Information criterion	8.60	Measure of the explanatory capability of the model proposed by Akaike (1974)
Schwarz Information criterion	8.70	Measure of the explanatory capability of the model following a Bayesian approach suggested by Schwarz (1978)
Durbin-Watson statistic	2.27	Measure of the first-order autocorrelation

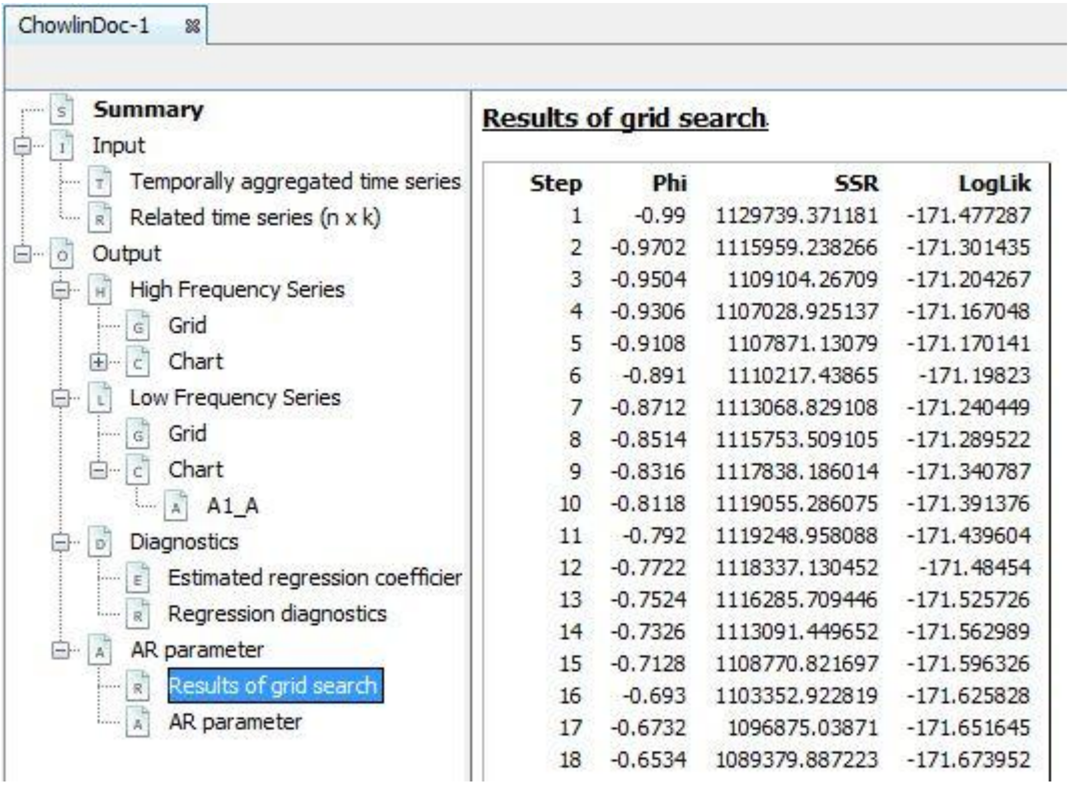
As can be deduced, all these results together with the "Output of the scanning procedure", which gives the estimated value for the first-order autoregressive parameter, are the most relevant output for the Chow-Lin technique.

Figure 47. Estimate AR parameter (Chow-Lin's).



Finally, the JEcotrim also returns the entire output from the scanning procedure.

Figure 48. Scanning procedure output (Chow-Lin's).

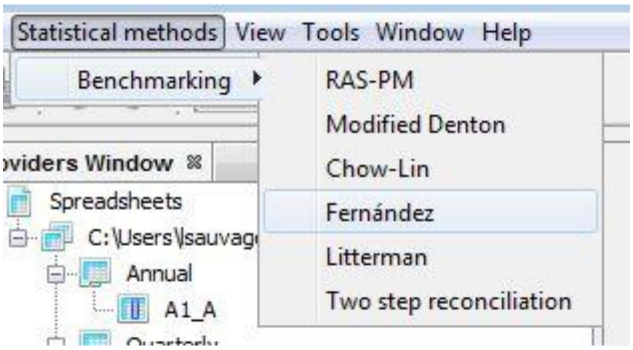


The other methods are following the same pattern in terms of presentation, of way to interact with the method and availability of output.

3.1.2. FERNÁNDEZ (1981)

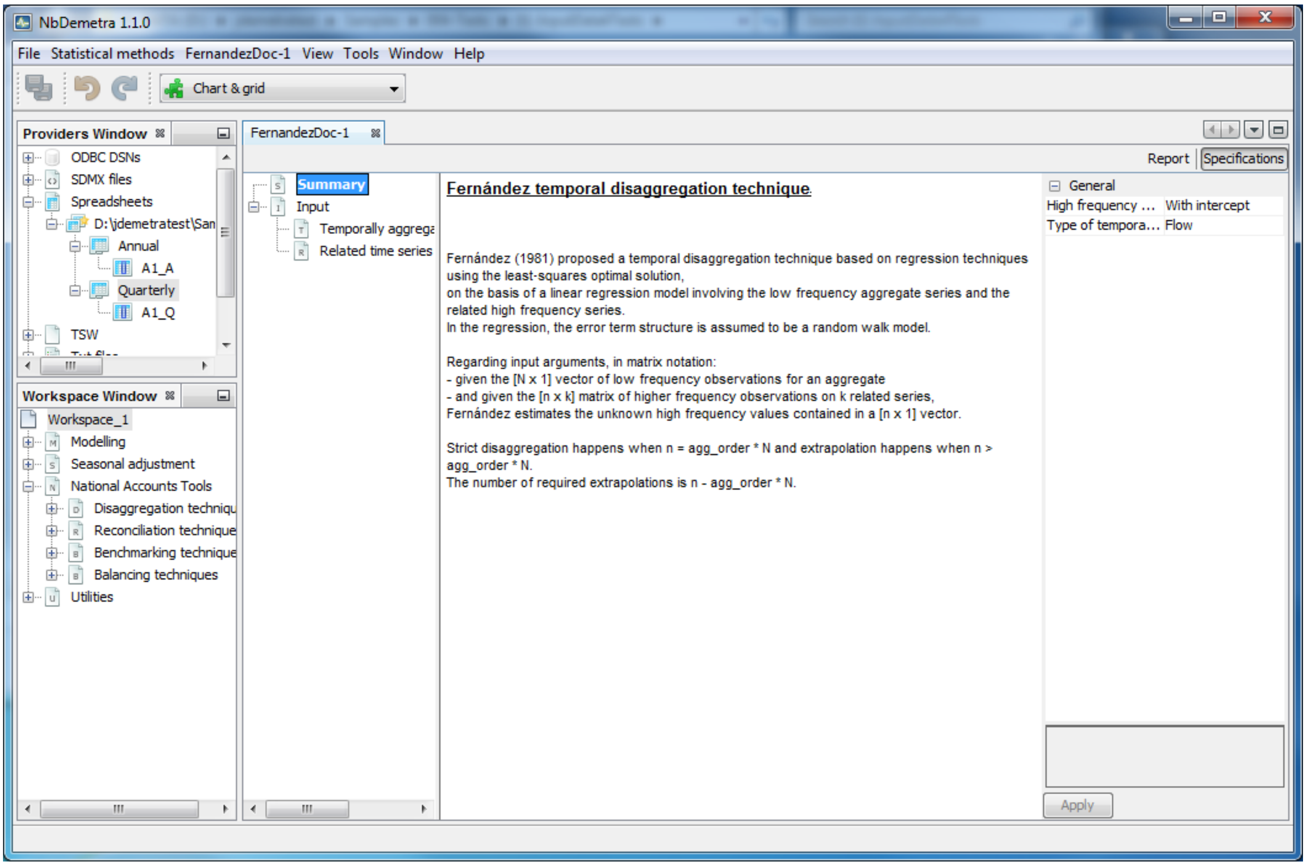
This method differs from the Chow-Lin's in the hypothesis related to the structure of the error in the regression model. In this case, the disturbance follows a random walk model. To use the Fernández merhod, select Statistical methods>National Accounts Tools>Fernández from the pull-down menu as can be seen in Figure 49.

Figure 49. Fernández technique selection.



Once the method is chosen, Figure 50 presents the same aspect as in Chow-Lin's: left-hand side for the "Summary", "Input", etc.; the middle, for a detailed description of the technique and the right-hand side brings the "Specifications" menu.

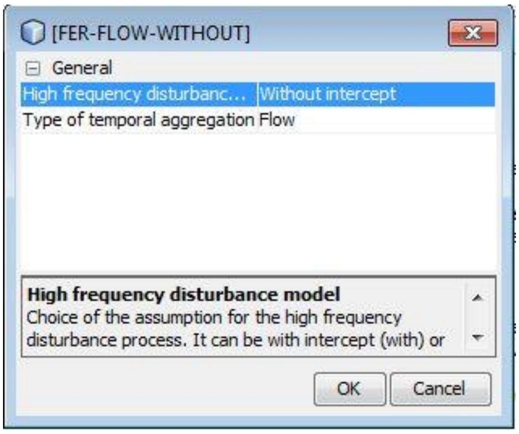
Figure 50. Overview of the main window for Fernández's.



Once the series have been loaded correctly, Fernández technique only permits the user to make the following modifications in the "Specifications" menu:

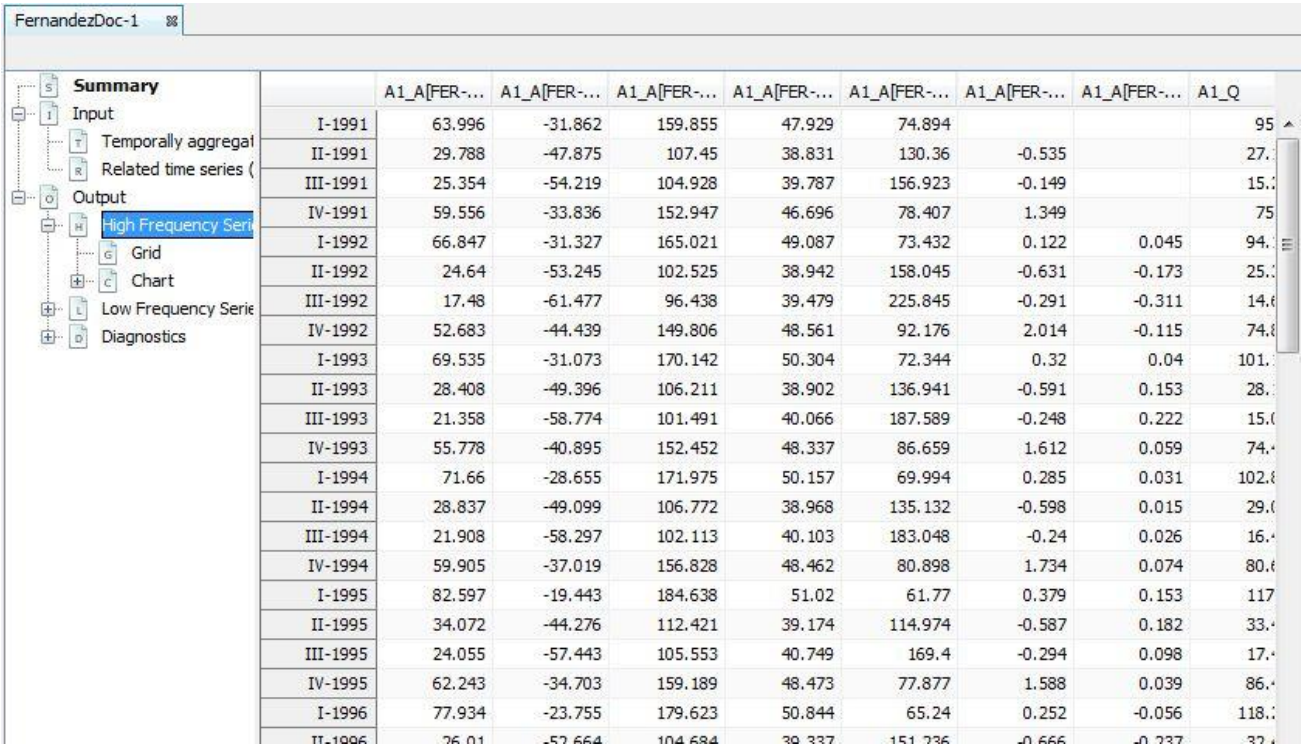
- whether a constant term is included or not in the high-frequency model, and
- the type of the temporal aggregation

Figure 51. Options for Fernández technique.



After this set-up stage, the interface returns the output. In this case, the "Output" menu comprises the "Grid" (Figure 52) and the "Chart" for the estimated series. The latter is not presented in the manual in order to avoid excessive repetition.

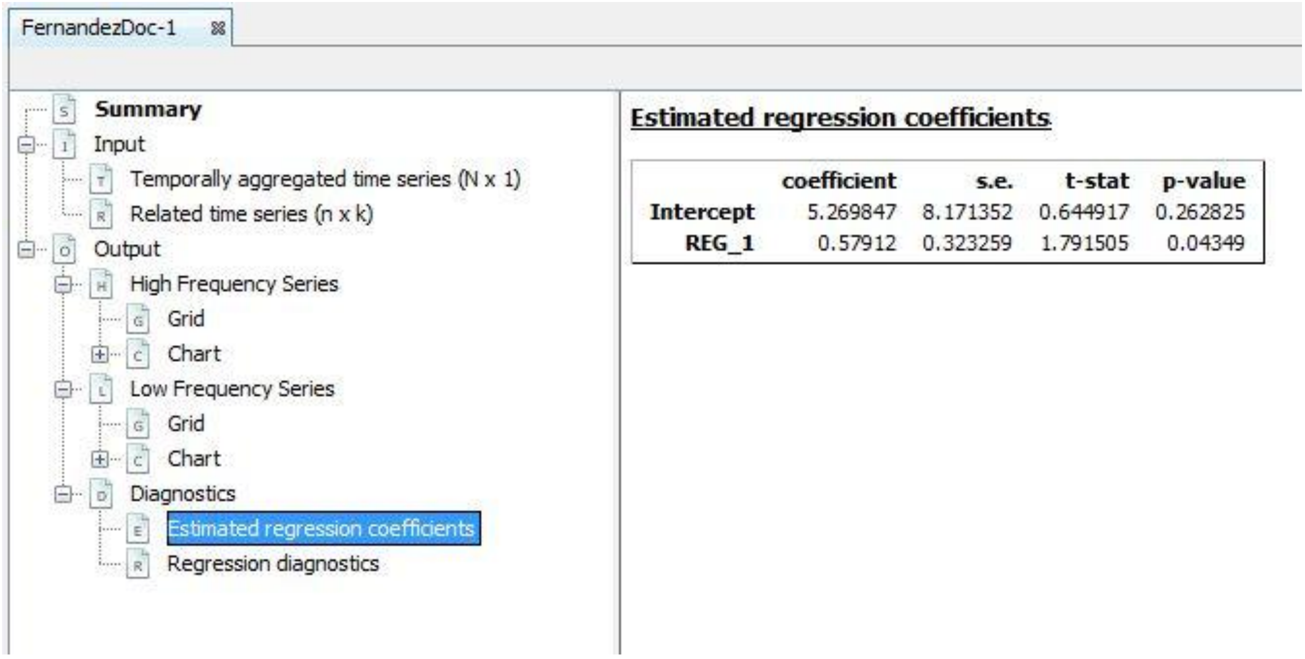
Figure 52. Grid of the disaggregated series (Fernández's).



	A1_A[FER-...	A1_A[FER-...	A1_A[FER-...	A1_A[FER-...	A1_A[FER-...	A1_A[FER-...	A1_A[FER-...	A1_Q
I-1991	63.996	-31.862	159.855	47.929	74.894			95
II-1991	29.788	-47.875	107.45	38.831	130.36	-0.535		27.
III-1991	25.354	-54.219	104.928	39.787	156.923	-0.149		15.
IV-1991	59.556	-33.836	152.947	46.696	78.407	1.349		75
I-1992	66.847	-31.327	165.021	49.087	73.432	0.122	0.045	94.
II-1992	24.64	-53.245	102.525	38.942	158.045	-0.631	-0.173	25.
III-1992	17.48	-61.477	96.438	39.479	225.845	-0.291	-0.311	14.
IV-1992	52.683	-44.439	149.806	48.561	92.176	2.014	-0.115	74.
I-1993	69.535	-31.073	170.142	50.304	72.344	0.32	0.04	101.
II-1993	28.408	-49.396	106.211	38.902	136.941	-0.591	0.153	28.
III-1993	21.358	-58.774	101.491	40.066	187.589	-0.248	0.222	15.
IV-1993	55.778	-40.895	152.452	48.337	86.659	1.612	0.059	74.
I-1994	71.66	-28.655	171.975	50.157	69.994	0.285	0.031	102.
II-1994	28.837	-49.099	106.772	38.968	135.132	-0.598	0.015	29.
III-1994	21.908	-58.297	102.113	40.103	183.048	-0.24	0.026	16.
IV-1994	59.905	-37.019	156.828	48.462	80.898	1.734	0.074	80.
I-1995	82.597	-19.443	184.638	51.02	61.77	0.379	0.153	117
II-1995	34.072	-44.276	112.421	39.174	114.974	-0.587	0.182	33.
III-1995	24.055	-57.443	105.553	40.749	169.4	-0.294	0.098	17.
IV-1995	62.243	-34.703	159.189	48.473	77.877	1.588	0.039	86.
I-1996	77.934	-23.755	179.623	50.844	65.24	0.252	-0.056	118.
II-1996	26.01	-52.664	104.684	39.337	151.236	-0.666	-0.237	32.

Furthermore, the obtained estimates can be found in Diagnostics>Estimated regression coefficients.

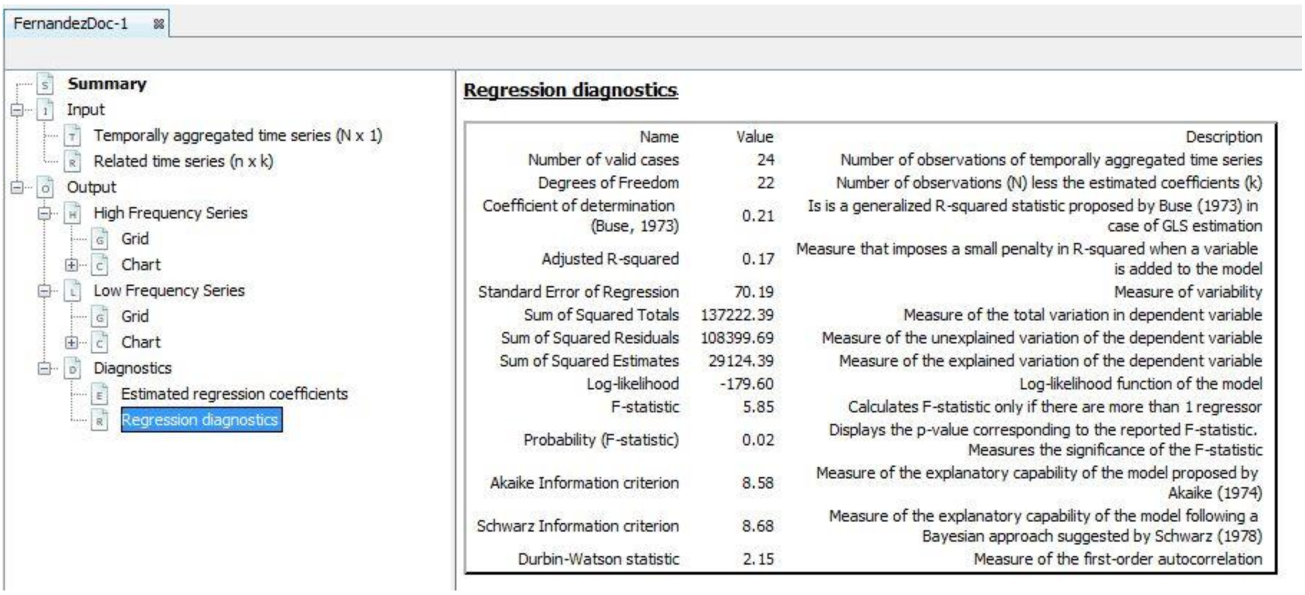
Figure 53. Estimated regression coefficients (Fernández's).



	coefficient	s.e.	t-stat	p-value
Intercept	5.269847	8.171352	0.644917	0.262825
REG_1	0.57912	0.323259	1.791505	0.04349

Moreover, Figure 54 shows the summary of the statistics to assess the accuracy of the estimates.

Figure 54. Summary of statistics (Fernández's).



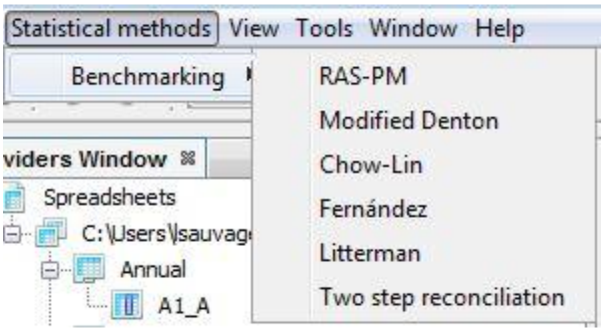
Name	Value	Description
Number of valid cases	24	Number of observations of temporally aggregated time series
Degrees of Freedom	22	Number of observations (N) less the estimated coefficients (k)
Coefficient of determination (Buse, 1973)	0.21	Is is a generalized R-squared statistic proposed by Buse (1973) in case of GLS estimation
Adjusted R-squared	0.17	Measure that imposes a small penalty in R-squared when a variable is added to the model
Standard Error of Regression	70.19	Measure of variability
Sum of Squared Totals	137222.39	Measure of the total variation in dependent variable
Sum of Squared Residuals	108399.69	Measure of the unexplained variation of the dependent variable
Sum of Squared Estimates	29124.39	Measure of the explained variation of the dependent variable
Log-likelihood	-179.60	Log-likelihood function of the model
F-statistic	5.85	Calculates F-statistic only if there are more than 1 regressor
Probability (F-statistic)	0.02	Displays the p-value corresponding to the reported F-statistic. Measures the significance of the F-statistic
Akaike Information criterion	8.58	Measure of the explanatory capability of the model proposed by Akaike (1974)
Schwarz Information criterion	8.68	Measure of the explanatory capability of the model following a Bayesian approach suggested by Schwarz (1978)
Durbin-Watson statistic	2.15	Measure of the first-order autocorrelation

Note that in this case, there is no parameter of interest since the model assumed for the error is a pure random walk model.

3.1.3. LITTERMAN (1983)

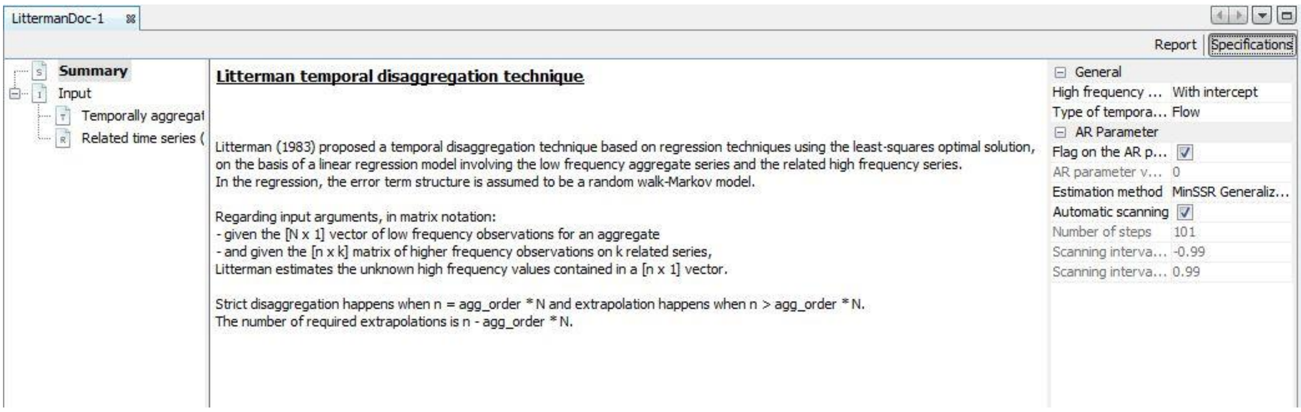
In this case, Litterman replaces the error structure that assumes Chow-Lin's by an ARIMA (0,1,1). The rest of the procedure remains unchanged. To use the Fernández merhod, select Statistical methods>National Accounts Tools>Fernández from the pull-down menu as can be seen in the figure below.

Figure 55. Litterman method selection.



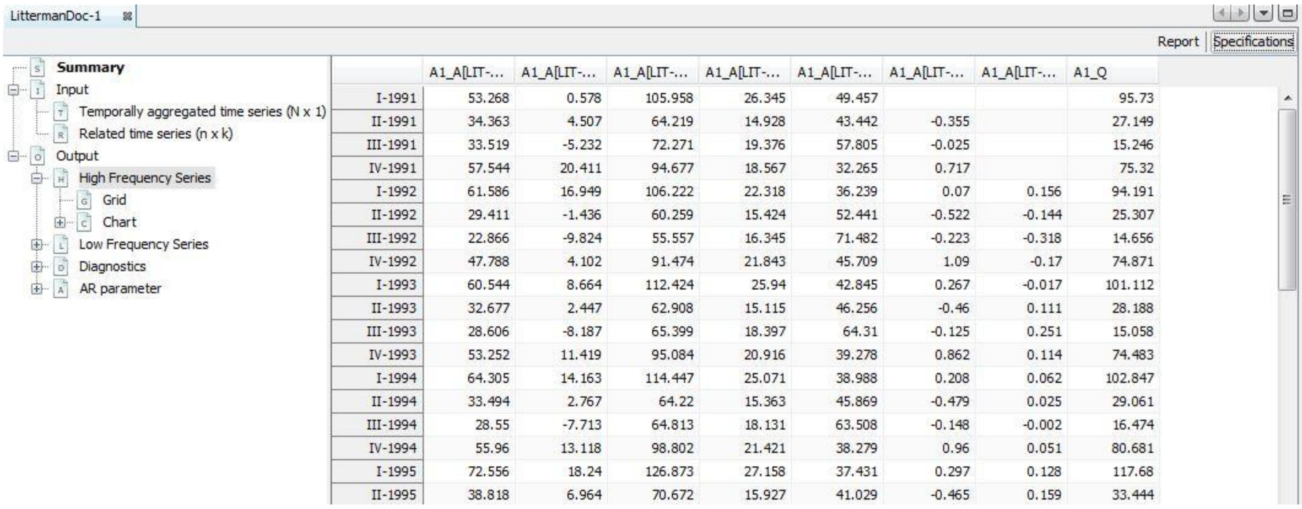
Next, we present the front window for this technique. Notice that the structure is similar to that explained in Chow and Lin (1971): the "Summary"; "Input", etc on the left side; the middle for the description of the method and the right side goes for the "Specifications" menu.

Figure 56. Main window for Litterman method (overview).



Once the series are loaded as explained before, this disaggregation method allows the user to modify all of the options the Chow-Lin method permitted (from Figure 35 to Figure 41). Then, after having made the proper modifications, the program yields:

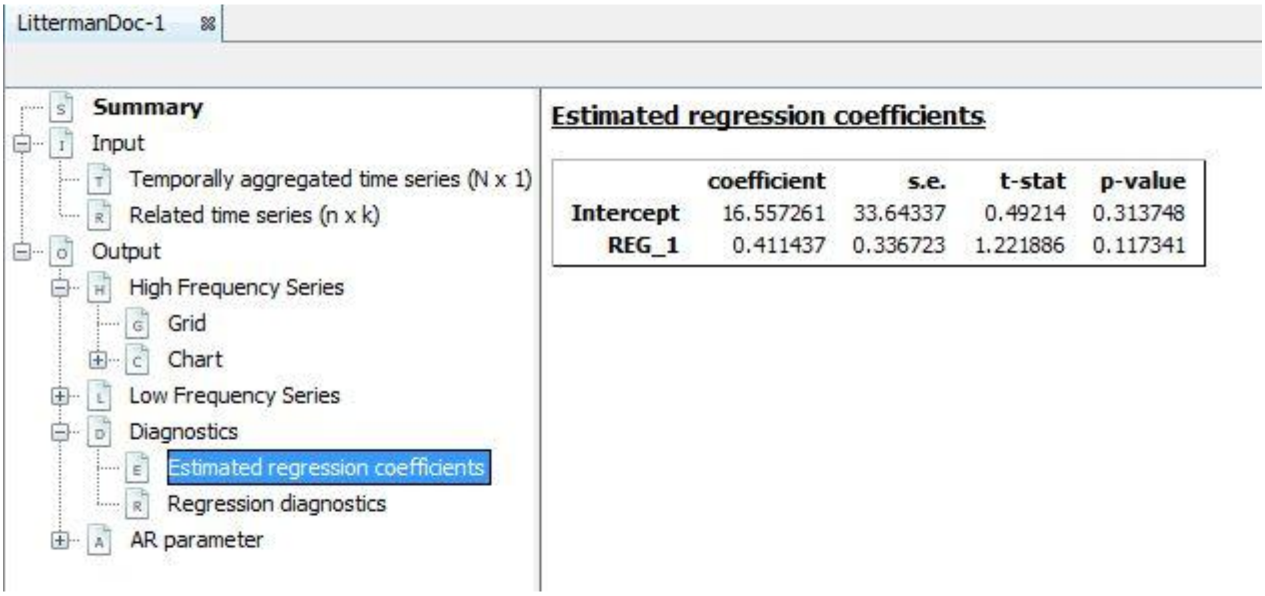
Figure 57. Grid of the disaggregated series (Litterman's).



	A1_A[LIT-...	A1_A[LIT-...	A1_A[LIT-...	A1_A[LIT-...	A1_A[LIT-...	A1_A[LIT-...	A1_A[LIT-...	A1_Q
I-1991	53.268	0.578	105.958	26.345	49.457			95.73
II-1991	34.363	4.507	64.219	14.928	43.442	-0.355		27.149
III-1991	33.519	-5.232	72.271	19.376	57.805	-0.025		15.246
IV-1991	57.544	20.411	94.677	18.567	32.265	0.717		75.32
I-1992	61.586	16.949	106.222	22.318	36.239	0.07	0.156	94.191
II-1992	29.411	-1.436	60.259	15.424	52.441	-0.522	-0.144	25.307
III-1992	22.866	-9.824	55.557	16.345	71.482	-0.223	-0.318	14.656
IV-1992	47.788	4.102	91.474	21.843	45.709	1.09	-0.17	74.871
I-1993	60.544	8.664	112.424	25.94	42.845	0.267	-0.017	101.112
II-1993	32.677	2.447	62.908	15.115	46.256	-0.46	0.111	28.188
III-1993	28.606	-8.187	65.399	18.397	64.31	-0.125	0.251	15.058
IV-1993	53.252	11.419	95.084	20.916	39.278	0.862	0.114	74.483
I-1994	64.305	14.163	114.447	25.071	38.988	0.208	0.062	102.847
II-1994	33.494	2.767	64.22	15.363	45.869	-0.479	0.025	29.061
III-1994	28.55	-7.713	64.813	18.131	63.508	-0.148	-0.002	16.474
IV-1994	55.96	13.118	98.802	21.421	38.279	0.96	0.051	80.681
I-1995	72.556	18.24	126.873	27.158	37.431	0.297	0.128	117.68
II-1995	38.818	6.964	70.672	15.927	41.029	-0.465	0.159	33.444

Also, the estimated regression coefficients within the "Diagnostics" branch:

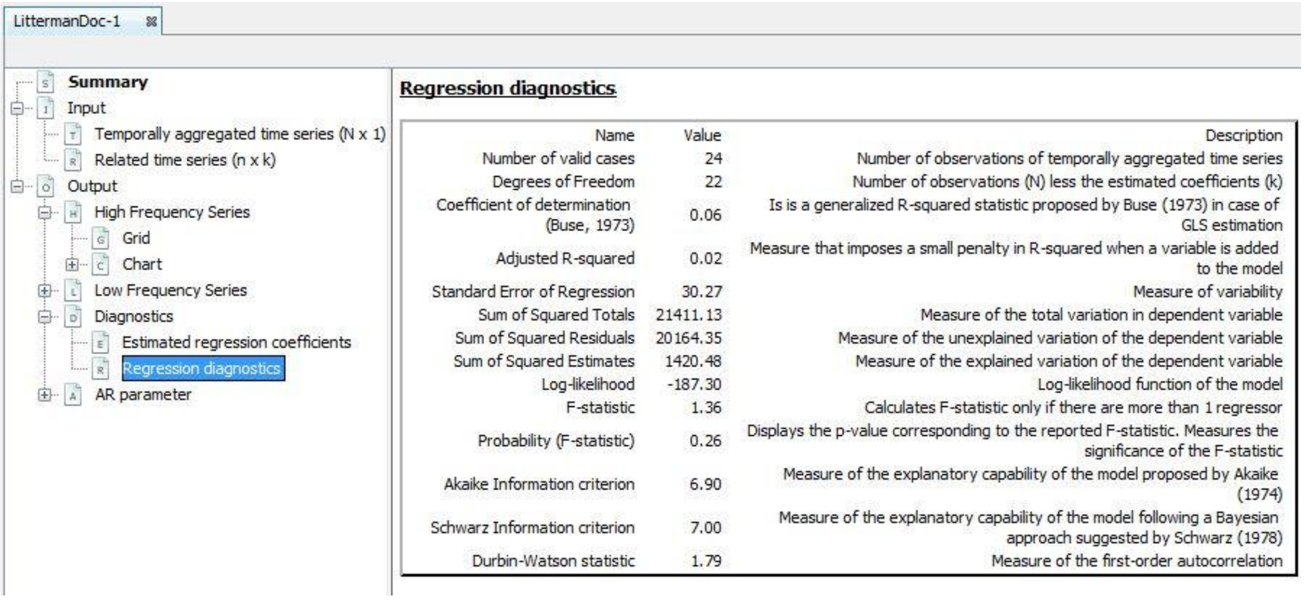
Figure 58. Estimated regression coefficients (Litterman's).



	coefficient	s.e.	t-stat	p-value
Intercept	16.557261	33.64337	0.49214	0.313748
REG_1	0.411437	0.336723	1.221886	0.117341

In order to evaluate the goodness of the estimates, JEcotrim provides the summary of the different statistics as can be seen next.

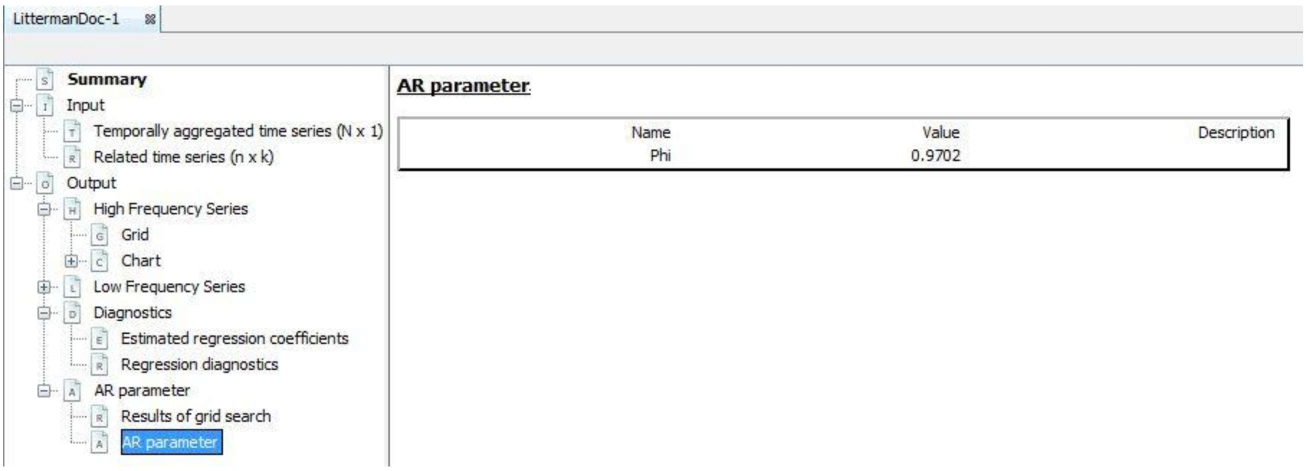
Figure 59. Summary of statistics (Litterman's).



Name	Value	Description
Number of valid cases	24	Number of observations of temporally aggregated time series
Degrees of Freedom	22	Number of observations (N) less the estimated coefficients (k)
Coefficient of determination (Buse, 1973)	0.06	Is is a generalized R-squared statistic proposed by Buse (1973) in case of GLS estimation
Adjusted R-squared	0.02	Measure that imposes a small penalty in R-squared when a variable is added to the model
Standard Error of Regression	30.27	Measure of variability
Sum of Squared Totals	21411.13	Measure of the total variation in dependent variable
Sum of Squared Residuals	20164.35	Measure of the unexplained variation of the dependent variable
Sum of Squared Estimates	1420.48	Measure of the explained variation of the dependent variable
Log-likelihood	-187.30	Log-likelihood function of the model
F-statistic	1.36	Calculates F-statistic only if there are more than 1 regressor
Probability (F-statistic)	0.26	Displays the p-value corresponding to the reported F-statistic. Measures the significance of the F-statistic
Akaike Information criterion	6.90	Measure of the explanatory capability of the model proposed by Akaike (1974)
Schwarz Information criterion	7.00	Measure of the explanatory capability of the model following a Bayesian approach suggested by Schwarz (1978)
Durbin-Watson statistic	1.79	Measure of the first-order autocorrelation

Also, the user can find the estimate value of the parameter of interest of the regression model:

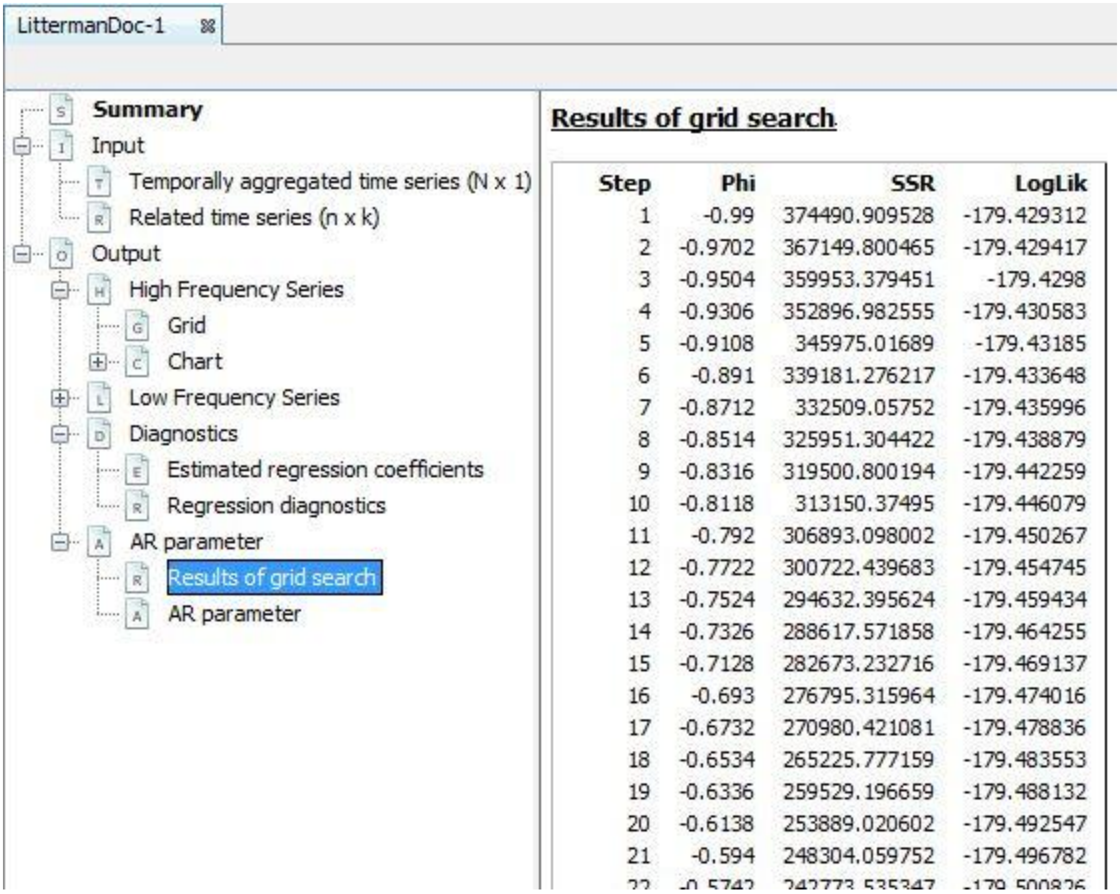
Figure 60. Estimate value for the parameter of interest.



Name	Value	Description
Phi	0.9702	

And, last but not least, one can discover the complete output for the scanning procedure as in Chow-Lin's.

Figure 61. Output of the scanning procedure (Litterman's).



Step	Phi	SSR	LogLik
1	-0.99	374490.909528	-179.429312
2	-0.9702	367149.800465	-179.429417
3	-0.9504	359953.379451	-179.4298
4	-0.9306	352896.982555	-179.430583
5	-0.9108	345975.01689	-179.43185
6	-0.891	339181.276217	-179.433648
7	-0.8712	332509.05752	-179.435996
8	-0.8514	325951.304422	-179.438879
9	-0.8316	319500.800194	-179.442259
10	-0.8118	313150.37495	-179.446079
11	-0.792	306893.098002	-179.450267
12	-0.7722	300722.439683	-179.454745
13	-0.7524	294632.395624	-179.459434
14	-0.7326	288617.571858	-179.464255
15	-0.7128	282673.232716	-179.469137
16	-0.693	276795.315964	-179.474016
17	-0.6732	270980.421081	-179.478836
18	-0.6534	265225.777159	-179.483553
19	-0.6336	259529.196659	-179.488132
20	-0.6138	253889.020602	-179.492547
21	-0.594	248304.059752	-179.496782

3.2. BENCHMARKING TECHNIQUE – Modified Denton (1984)

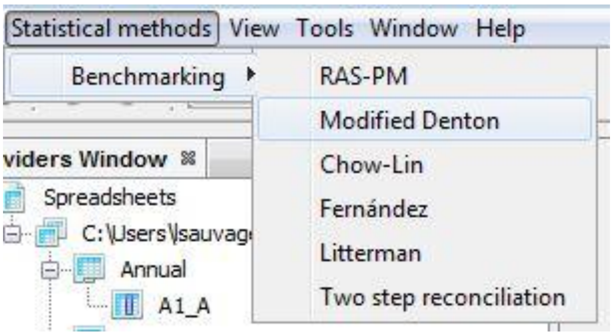
This section considers the proposal of Cholette (1984) to modify the solution originally proposed by Denton (1971). This version differs from the seminal work with respect to the treatment of the starting condition. So as not to enter into so much detail, we can say that this variation consists, ultimately, in a constrained minimization problem in which an aggregated series and a preliminary series are involved and that returns the corresponding estimates while the original movements of the preliminary series are preserved.

Regarding to these movements we can distinguish different options to be implemented in the previous technique: additive first differences (AFD), additive second differences (ASD), proportional first differences (PFD) and proportional second differences (PSD) (see Annex 4). As the user will see further, the JEcotrim interface collects all these movements.

In this guide, we will present only the results for PFD which, from our point of view, results to be the most adequate solution in many cases. Furthermore, the procedure for using other techniques is very similar.

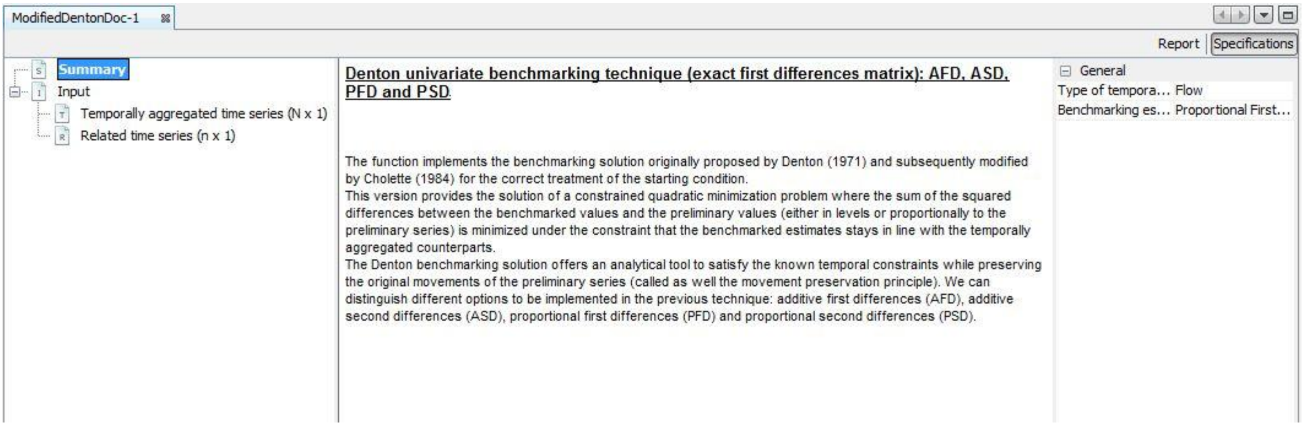
Thus, to select this method in JEcotrim the user should go to Statistical methods>National Accounts Tools>ModifiedDenton as shown in Figure 62

Figure 62. Modified Denton benchmarking technique selection.



The overview of the main window for this method is given by Figure 63[4]

Figure 63. Main window for Modified Denton technique (overview).

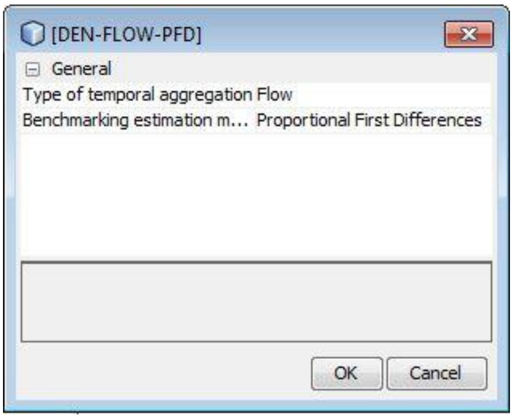


The first step is to load the series as previously described in this guide. Again, the series object of analysis is A1.xlsx. After that, the present method only allows the user to change the type of temporal aggregation (flow, index, first and last) and the benchmarking estimation method (AFD, ASD, PFD and PSD)[5]

[4] Notice that the same structure as in the previous regression based methods is maintained.

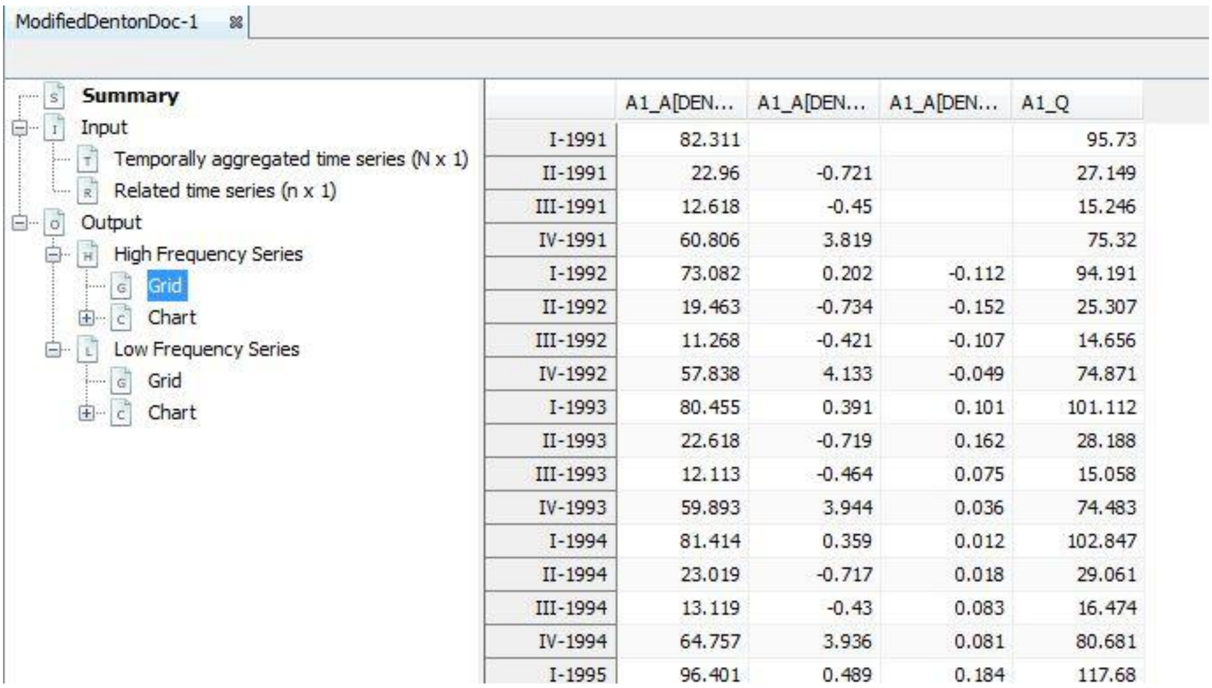
[5] See Annex 3.

Figure 64. Allowed modifications in the Modified Denton technique.



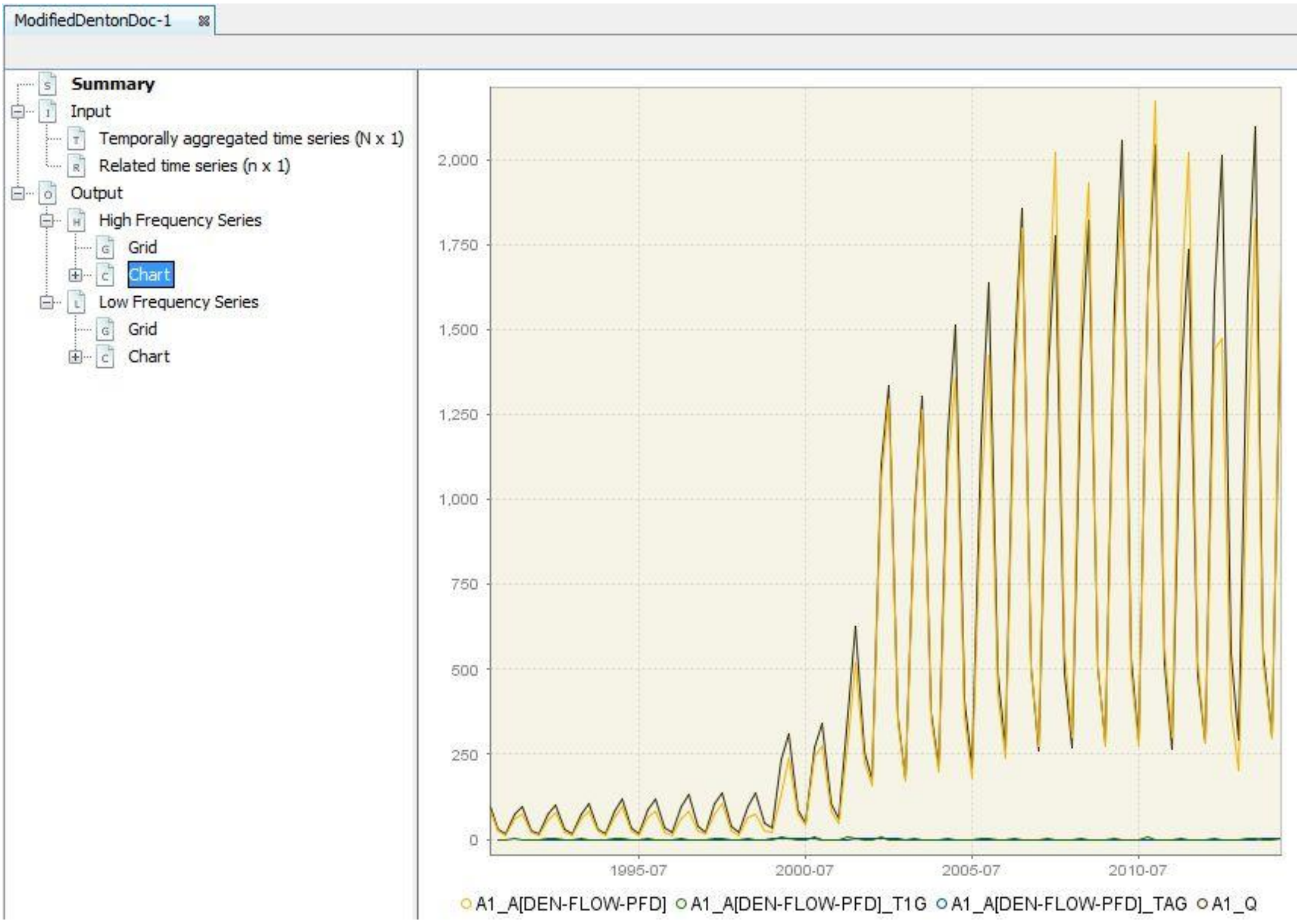
After having made the corresponding variations, the interface yields the output. Here one can find the "Grid" and the "Chart" of the benchmarked series (1st column), its lag1 growth rate (2nd column), its annual growth rate (3rd column) and the original related series (4th column), as can be seen in Figure 65 and Figure 66.

Figure 65. Benchmarked series.



	A1_A[DEN...	A1_A[DEN...	A1_A[DEN...	A1_Q
I-1991	82.311			95.73
II-1991	22.96	-0.721		27.149
III-1991	12.618	-0.45		15.246
IV-1991	60.806	3.819		75.32
I-1992	73.082	0.202	-0.112	94.191
II-1992	19.463	-0.734	-0.152	25.307
III-1992	11.268	-0.421	-0.107	14.656
IV-1992	57.838	4.133	-0.049	74.871
I-1993	80.455	0.391	0.101	101.112
II-1993	22.618	-0.719	0.162	28.188
III-1993	12.113	-0.464	0.075	15.058
IV-1993	59.893	3.944	0.036	74.483
I-1994	81.414	0.359	0.012	102.847
II-1994	23.019	-0.717	0.018	29.061
III-1994	13.119	-0.43	0.083	16.474
IV-1994	64.757	3.936	0.081	80.681
I-1995	96.401	0.489	0.184	117.68

Figure 66. Graphical representation of the benchmarked series.



3.3. RECONCILIATION TECHNIQUE – The Two step procedure

3.3.1. Principles of the algorithm

<u>Theoretical origin</u>

The implemented reconciliation technique is the 2-STEP procedure developed by Di Fonzo and Marini (2011). [6]There are two variants of the algorithms:

– Di Fonzo and Marini (2011), abbreviated as ST in the interface options,

[6] See bibliographical references

– Quenneville and Rancourt (2005), abbreviated as QR in the interface options.

JEcotrim includes both variants and the first one is the default.

Motivation

Why this method? The statistical domains often introduce two levels of constraints:

- the temporal one, where the high frequency (HF) data should be aligned on its corresponding low frequency (LF) data; for example, the sum of the quarters in the flow case, should be equal to the year;
- the contemporaneous one, where the declared aggregate should be equal to the sum[7] of its components.

Implementation in JEcotrim

The 2-STEP procedure is proposing an algorithm in two steps where:

- first, a univariate method is applied to get the alignment of HF data on LF data; in this implementation, Modified Denton PFD is applied at once on all series involved in the process, i.e. on the aggregate and its components;
- second, the Stone algorithm is applied to solve simultaneously the temporal and contemporaneous constraints starting with aligned data along the temporal dimension.

According to the publication as mentioned before, the simultaneous resolution is insured by unit of period along the low frequency data: for example, if LF data is yearly, the system is iterating along the dataset by blocks of yearly data one after the other, on the full time range; if the HF data is quarterly, the input dataset to be reconciled is a rectangle of 4 lines, because of 4 quarters and m columns if the aggregation formula contains m series.

An additional mechanism has been introduced if a year of HF data is not complete: for example, a current year may have 3 quarters available out of 4 and statisticians would like to compute them anyway. The 2-Step procedure is able to deal with such a dataset. It is treating the last loop (because the incomplete year of data is of course the last one), i.e. the last fully available year by concatenating the extra incomplete dataset of the last year. For example, if year 2012 is complete and year 2013 has only 3 quarters available, the last loop will be a dataset of 7 lines for as columns as there are components. It will be called the incomplete-year mechanism.

As the series may enter in the 2-STEP procedure already aligned along the temporal constraint, a threshold has been implemented to decide automatically if it is the case (see the 'First step tolerance' in the interface parameters). If the aggregation of all HF data inside a unit of LF data is under the threshold, the system skips the Modified Denton PFD. Consequently, the user may solve the temporal constraint through a preliminary computation of all involved series in other univariate methods of the software, i.e. any regression-based method.

---

[7] CAUTION: The sum of the components may be a weighed one.

3.3.2. Usage from the interface

The ST method is selected as follows

Figure 67. Two-step procedure selection.

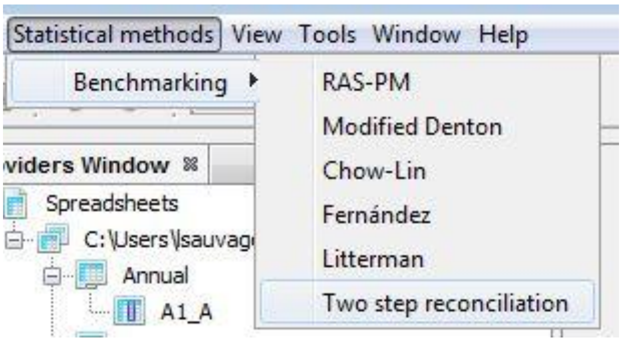
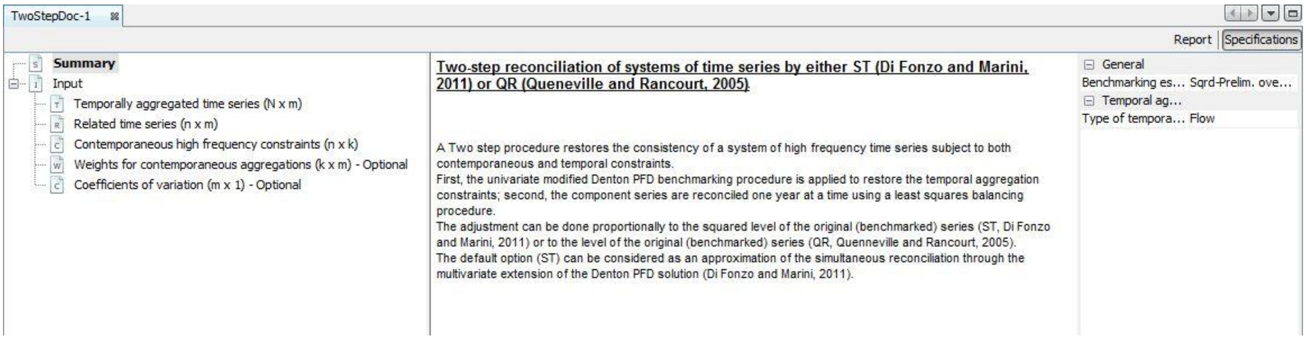


Figure 68 contains the overview of the main window for the corresponding method[8].

Figure 68. Main window for Two-step method.



After having selected the Two-step procedure, the user has to load the dataset[9] as explained before. Once the dataset is loaded, then in order to activate the input series the user has to follow for each input series the action shown in Figure 69 (drag and drop). This action should be repeated at least three times for the 3 mandatory inputs (Temporally aggregated times series, Related time series and Contemporaneous high frequency constraints).

Note that this method introduced 2 specificities about the input in regard for all the methods described in the previous chapters:

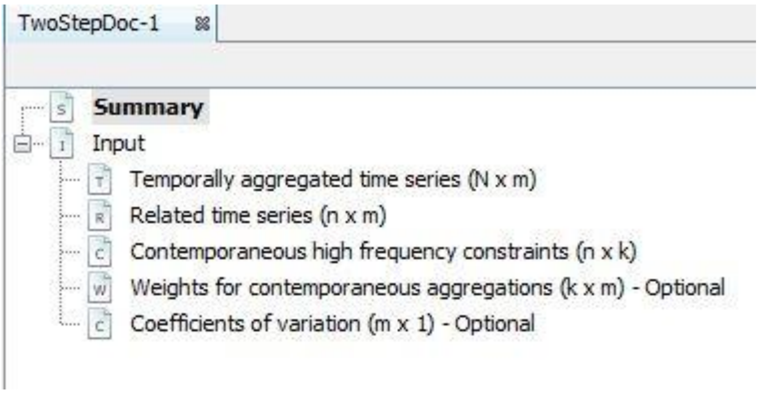
---

[8] Again, the present window follows the same structure as in previous methods.

[9] In this case, the dataset employed is called 2StepY1 and 2StepY2.

– The presence of optional input data. These two optional inputs (Weights for contemporaneous aggregations and Coefficients of variation) are not required by JEcotrim to execute the Two-step procedure. When not specified by the user, these parameters take default values, as described below;
– These two optional parameters are not time series but tables and thus have to be loaded by a different procedure, not using the JDemetra+ providers. This procedure is described below.

Figure 69. Inputs in a Two-step procedure.

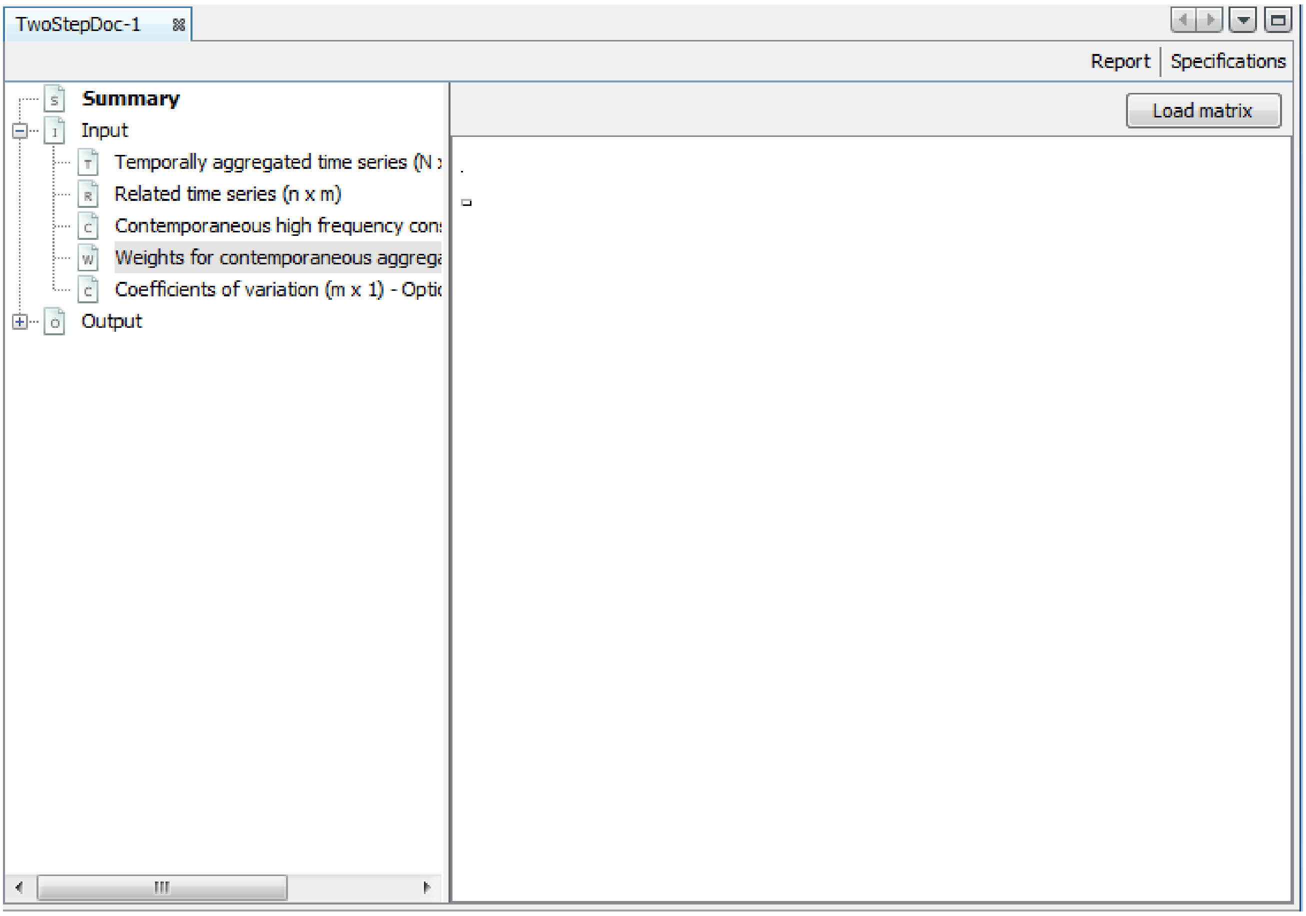


The optional input named "Weights for contemporaneous aggregations" is a matrix that defines the contemporaneous linkages among the y's at the high-frequency. The default value is a k x m matrix full of ones, where k is the number of contemporaneous constraints and m the number of temporally aggregated time series in the system studied.

The optional input named "Coefficients of variation" has to contain the coefficients of variation for different variables reliability. By default, it is a column vector of m values all equal to 1 where m is the number of temporally aggregated time series in the system studied.

As you can see in the picture below, these 2 optional inputs present a button named "Load matrix" in the upper-right part of the window, meaning that they cannot be loaded through a classic drag-and-drop from the providers as before, because they are not composed of time series. Instead, the user has to click on the button "Load matrix".

Figure 70. The "Load matrix" button.

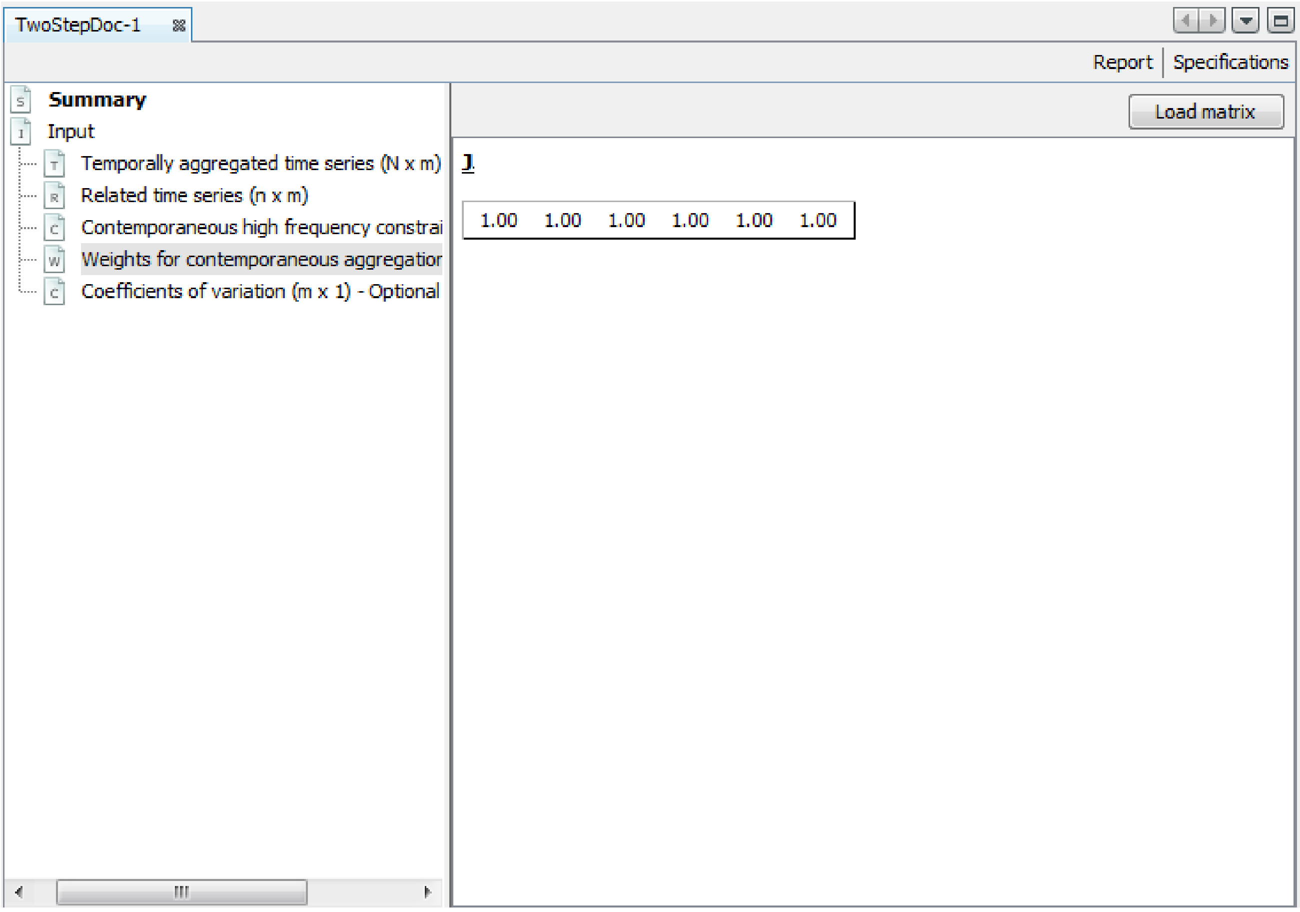


Clicking on this button launches a classic "Open" window, where you can browse your file system to locate the file containing the matrix you wish to load.

JEcotrim will look for the matrix in the first sheet of the workbook only and the table must begin in cell A1 of this sheet, without any headers.

The figure below shows what the window looks like after loading a matrix (here a row vector of 6 values, all equal to 1).

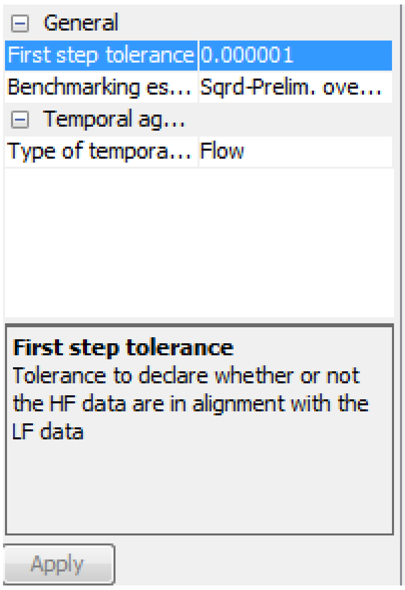
Figure 71. Matrix loaded.



For the Two-step reconciliation, the "Specifications" menu has three options to be changed by the user:

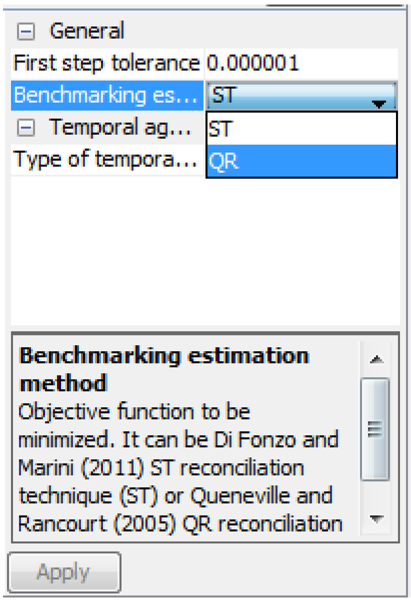
– The value of the First step tolerance.
  The tolerance is used in the preliminary check of the data, to determine if the 1st step should be applied or not.

Figure 72. Choice of the First step tolerance.



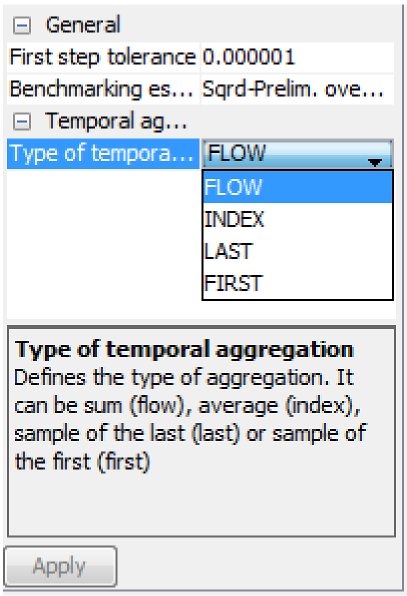
– The benchmarking method used for the reconciliation: ST (default) or QR.

Figure 73. Benchmarking estimation method selection.



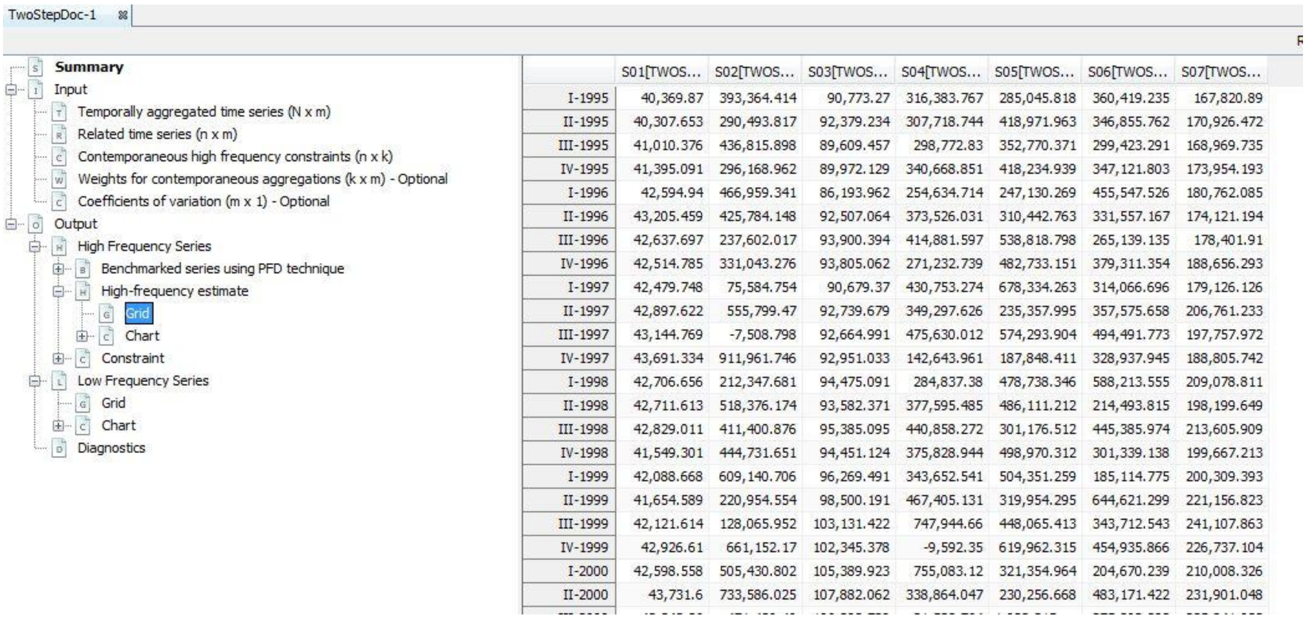
– The type of the temporal aggregation method used for the first and the second step: flow, index, first or last.

Figure 74. Type of temporal aggregation (Two-step method).



Once the user has made the corresponding choices, the JEcotrim interface yields the reconciliated series as shown in the “Grid” view just below

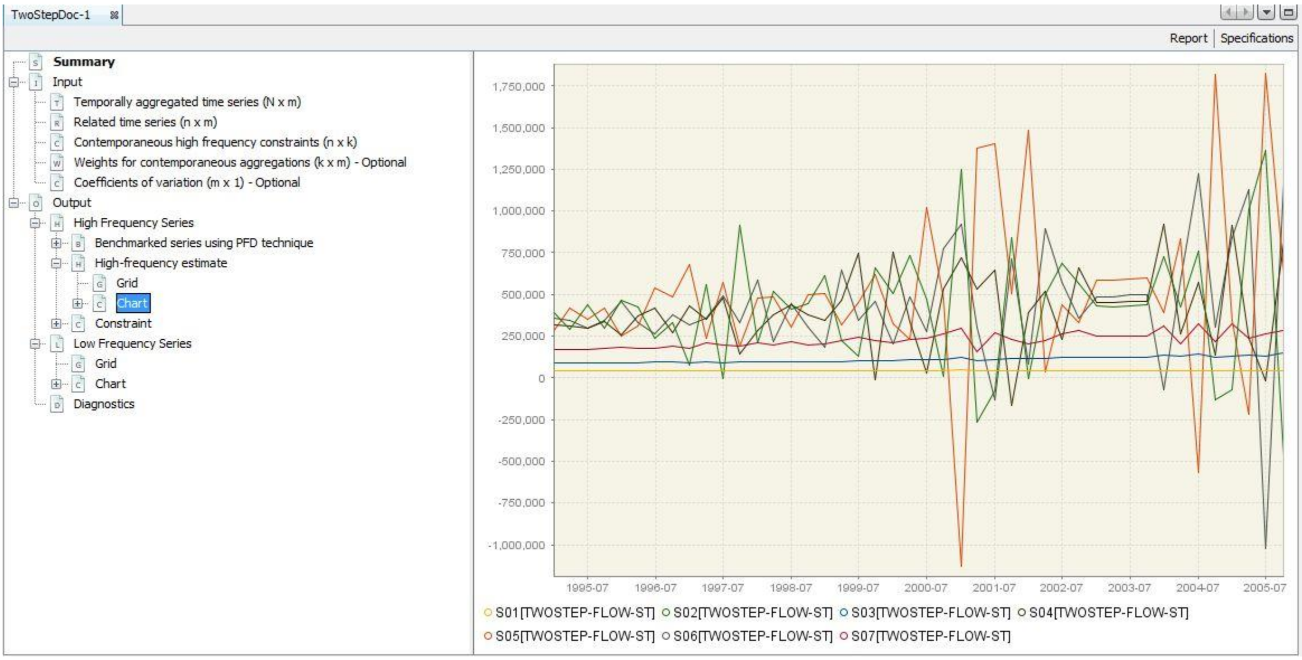
Figure 75. Grid of reconciliated series.



	S01[TWOS…	S02[TWOS…	S03[TWOS…	S04[TWOS…	S05[TWOS…	S06[TWOS…	S07[TWOS…
I-1995	40,369.87	393,364.414	90,773.27	316,383.767	285,045.818	360,419.235	167,820.89
II-1995	40,307.653	290,493.817	92,379.234	307,718.744	418,971.963	346,855.762	170,926.472
III-1995	41,010.376	436,815.898	89,609.457	298,772.83	352,770.371	299,423.291	168,969.735
IV-1995	41,395.091	296,168.962	89,972.129	340,668.851	418,234.939	347,121.803	173,954.193
I-1996	42,594.94	466,959.341	86,193.962	254,634.714	247,130.269	455,547.526	180,762.085
II-1996	43,205.459	425,784.148	92,507.064	373,526.031	310,442.763	331,557.167	174,121.194
III-1996	42,637.697	237,602.017	93,900.394	414,881.597	538,818.798	265,139.135	178,401.91
IV-1996	42,514.785	331,043.276	93,805.062	271,232.739	482,733.151	379,311.354	188,656.293
I-1997	42,479.748	75,584.754	90,679.37	430,753.274	678,334.263	314,066.696	179,126.126
II-1997	42,897.622	555,799.47	92,739.679	349,297.626	235,357.995	357,575.658	206,761.233
III-1997	43,144.769	-7,508.798	92,664.991	475,630.012	574,293.904	494,491.773	197,757.972
IV-1997	43,691.334	911,961.746	92,951.033	142,643.961	187,848.411	328,937.945	188,805.742
I-1998	42,706.656	212,347.681	94,475.091	284,837.38	478,738.346	588,213.555	209,078.811
II-1998	42,711.613	518,376.174	93,582.371	377,595.485	486,111.212	214,493.815	198,199.649
III-1998	42,829.011	411,400.876	95,385.095	440,858.272	301,176.512	445,385.974	213,605.909
IV-1998	41,549.301	444,731.651	94,451.124	375,828.944	498,970.312	301,339.138	199,667.213
I-1999	42,088.668	609,140.706	96,269.491	343,652.541	504,351.259	185,114.775	200,309.393
II-1999	41,654.589	220,954.554	98,500.191	467,405.131	319,954.295	644,621.299	221,156.823
III-1999	42,121.614	128,065.952	103,131.422	747,944.66	448,065.413	343,712.543	241,107.863
IV-1999	42,926.61	661,152.17	102,345.378	-9,592.35	619,962.315	454,935.866	226,737.104
I-2000	42,598.558	505,430.802	105,389.923	755,083.12	321,354.964	204,670.239	210,008.326
II-2000	43,731.6	733,586.025	107,882.062	338,864.047	230,256.668	483,171.422	231,901.048

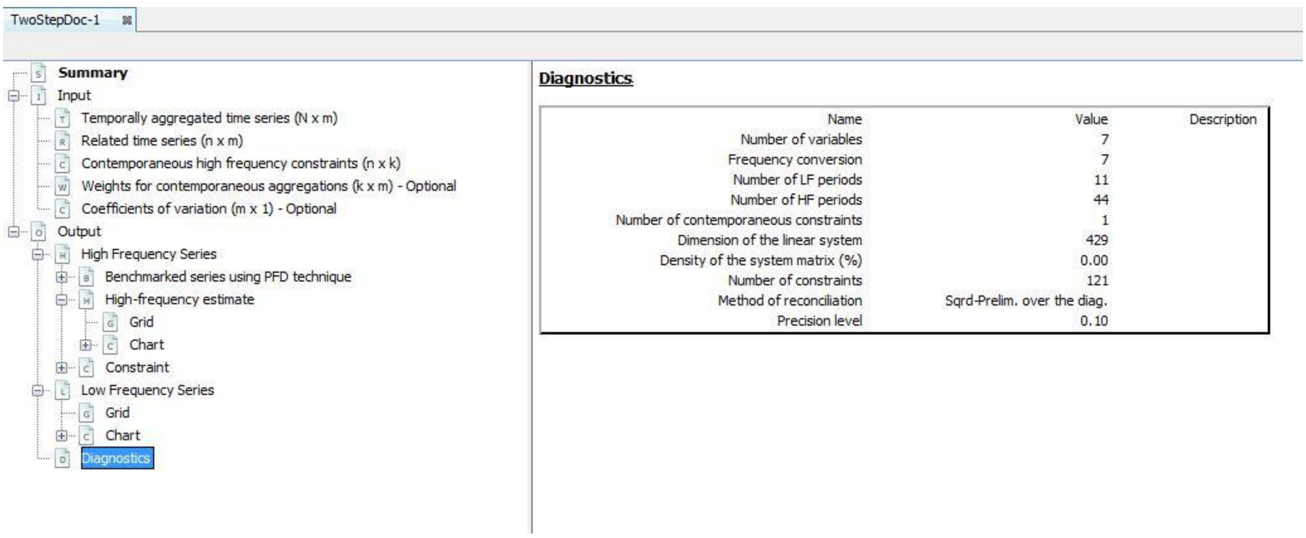
The “Chart” is also provided by the interface:

Figure 76. Chart of the reconciliated series.



Eventually, the “Output” drop down menu offers a summary of descriptive information (cf Annex 5):

Figure 77. Summary of descriptive information.



3.4. MATRIX BALANCING TECHNIQUE – RAS-PM (1964)

Matrix balancing is an important problem that has attracted attention in many different fields: the need for adjusting the entries of a large matrix to satisfy prior consistency requirements occurs frequently in economics, urban planning, statistics, demography, and stochastic modeling and a large amount of both theoretical results and real-life-data applications of matrix balancing can be found in the specialized literature for all these fields.

Thus, procedures for matrix balancing can be separated into two broad classes: bi-proportional or scaling algorithms (like RAS) and non bi-proportional algorithms. In this guide we focus our intention on the former, which are identified by the way they balance matrices[10].

The RAS algorithm is a bi-proportional matrix balancing procedure developed by Stone (1961) in an input-output context. Nevertheless, JEcotrim implements the Plus-Minus Proportionate Adjustment technique proposed by the Australian Bureau of Statistics in 1995[11]. This new method consists of a procedure for prorating a sequence of positive and negative values in line with a given total, while the original RAS algorithm can only deal with positive entries. Thus, the RAS-PM requires the use of two factors, one for the positive entries and one for the negative entries.

To use the RAS-PM in JEcotrim, the user has to select from Statistical methods>National Accounts Tools>RAS-PM

Figure 78. RAS-PM selection.

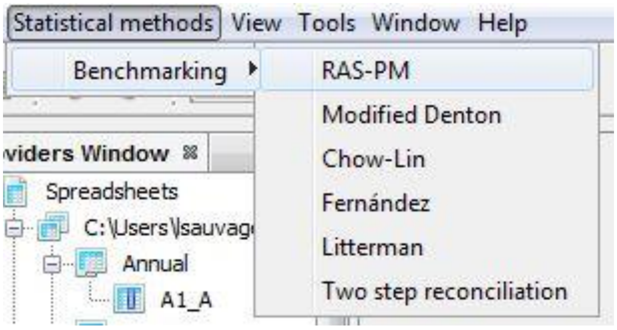


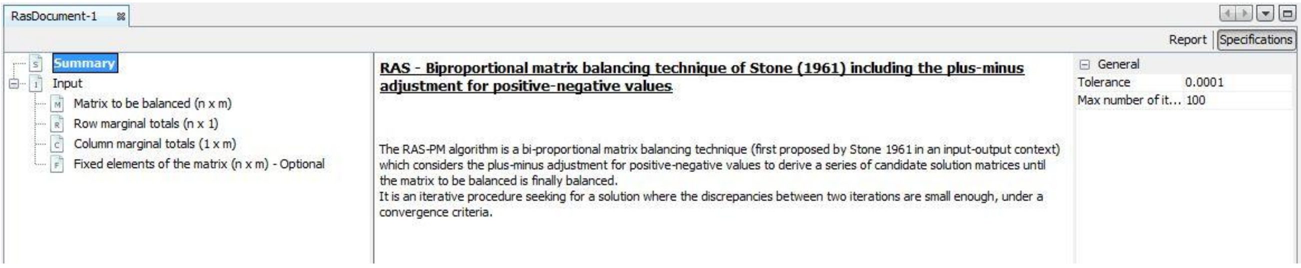
Figure 79 provides the overview of the main window of the Workspace. Again, the same structure as in previous cases is met.

---

[10] They iteratively multiply rows and columns of the matrix under adjustment by positive constants to derive a series of candidate solution matrices until the matrix is finally balanced.

[11] The work is called "Demographic estimates and projections: Concepts, Sources and Methods".

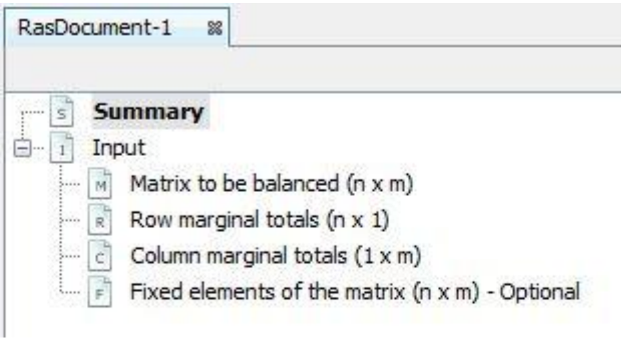
Figure 79. Main view of the RAS-PM Workspace (overview).



A word of caution is required in this case, because the user do not need load any dataset of time series as previously. Since this method works with matrices[12], the procedure to load them correctly is explained in the following lines (and also in the Two step procedure chapter).

Given the inputs in Figure 80

Figure 80. Inputs in RAS-PM technique.



First step is to select the first input “Matrix to be balanced” and then click on “Load matrix” button as indicated in Figure 81. This action should be repeated at least twice more[13].

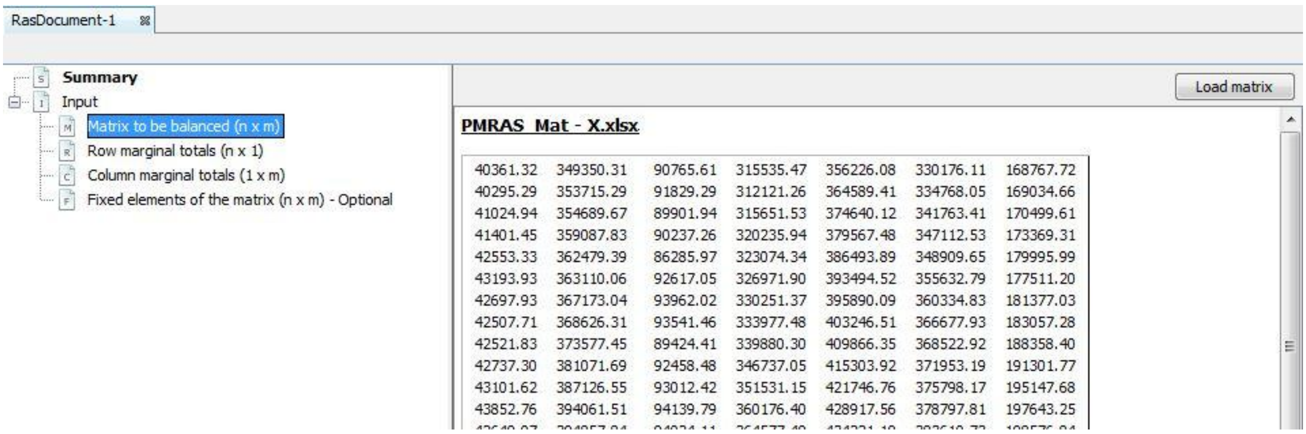
[12] The following Excel sheets are loaded: xmat.xlsx, colmarg.xlsx and rowmarg.xlsx.

[13] The last input called “Fixed elements of the matrix” is set default as a matrix of ones.

Figure 81. How to load matrices.

If the matrix to be balanced satisfies the conditions required by the procedure[14]

Figure 82. Loading matrix to be balanced.



And the rest of the input terms are correctly loaded (Figure 82 and Figure 83), then the program will not return any error message and will execute.

[14] See the Australian Bureau of Statistics' work for a depth view of the conditionals.

Figure 83. Loading row marginal totals.

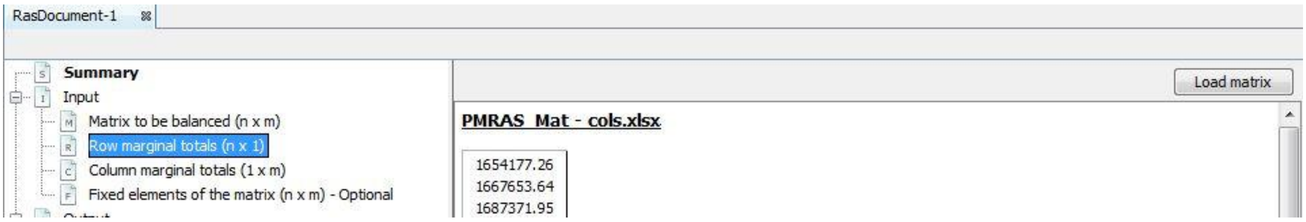
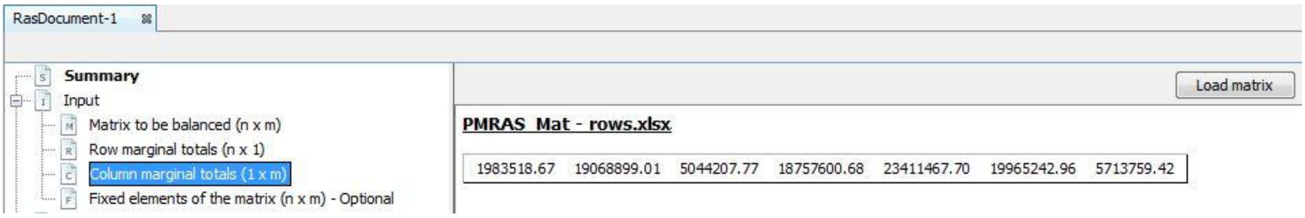


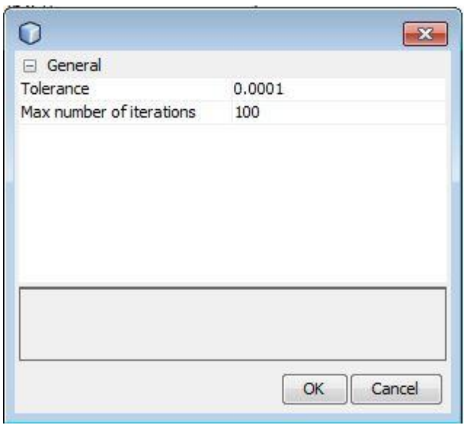
Figure 84. Loading column marginal totals.



Notice that we let the "Fixed elements of the matrix" being set by default.

JEcotrim only lets the user modify the level of tolerance for the convergence criterion which default value is 1,00E-04 and also the maximum number of iterations (100 is the default value in the interface).

Figure 85. Available modifications in "Specifications" menu for RAS-PM.

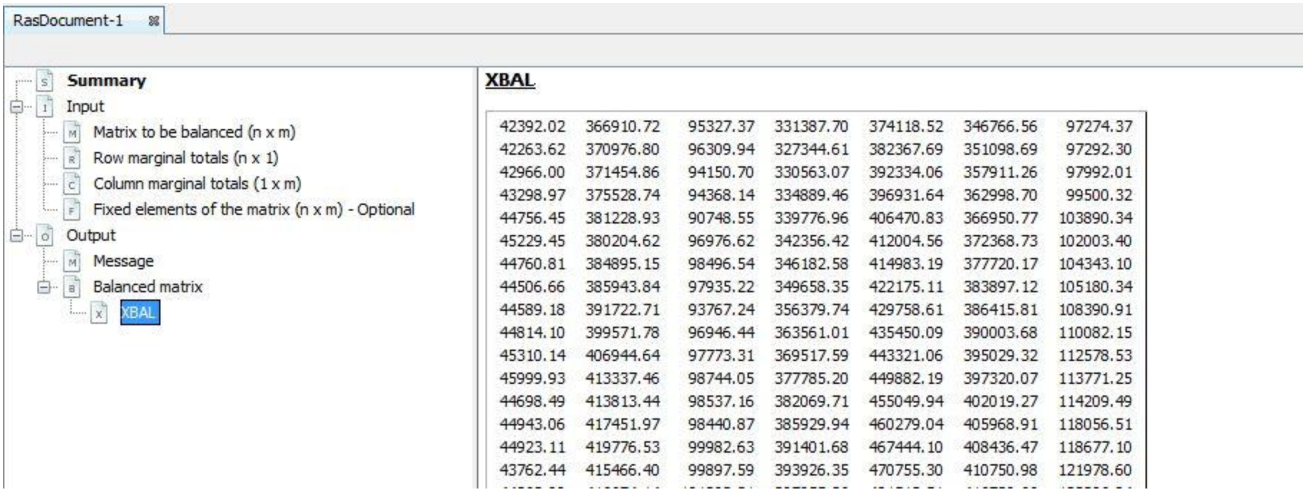


Once the set-up is established, the program runs automatically. First thing the user will run into in the "Output" menu is the message that informs whether the procedure has converged and the number of iterations required to that convergence.

Figure 86. Message of convergence for RAS-PM.

After the convergence message, the program presents the balanced matrix by RAS-PM, having the property to satisfy the required marginal totals.

Figure 87. Balanced matrix by RAS-PM.



4. Bridge from Ecotrim for Windows

This chapter gives indication to users of Ecotrim v.1.01 about what they can expect from JEcotrim and gives insight about how to go on in JEcotrim with the work they were doing in Ecotrim v.1.01.

This new program brings many improvements relative to the Ecotrim v.1.01. Compared to this last one, the JEcotrim package provides:

- User-friendly interface for loading data and output.
- Results for regression based techniques as for temporal disaggregation of time series have been stabilized when the GLS or ML method is employed for estimating. This is a remarkable point because in Ecotrim v.1.01 there seems to be a miscalculation of the ML method.
- The implemented univariate benchmarking method (Modified Denton) in JEcotrim returns quality and accurate output.
- A reconciliation procedure, namely Two-step reconciliation procedure (Di Fonzo and Marini 2011), has been implemented.
- Quality improvement of the results.
- A matrix balancing technique, namely the RAS-PM procedure, has been implemented.
- Neither Multivariate Denton nor Rossi techniques have been implemented since they seemed not to work in a proper way in Ecotrim v.1.01.

Regarding the theoretical methodology, there exist two popular approaches in the field of balancing/benchmarking time series: a Fully-Matrix Regression-Based (FMRB) approach using the least squares techniques, and a Recursive State-Space Regression Matrix (RSSRM) approach. The FMRB approach uses matrix calculation to provide exact results of parameters estimation while the RSSRM estimation results depend on the starting points of the iterative process. Such points are usually generated using diffuse distribution or similar techniques which may introduce a bias on the final results. However, the two approaches should provide identical results if the starting points are exact ones. Choosing one or the approach is a major factor that may affect the quality of the final results. Based on the large progress in the capabilities of computing technology, and anxious to produce exact results, the new generation of ECOTRIM uses the FMRB as the basic approach.

ECOTRIM offers the methods as in the table below.

- ***Univariate methods in ECOTRIM***

The univariate methods option is selected in order to estimate high-frequency univariate series. Ecotrim proposes either adjustment or optimal techniques, in the least squares sense. This module caters for two categories of methods of disaggregation:

a. Mathematical methods, which do not use reference indicators.

b. Statistical methods, which do use reference indicators.

Table 1: Methods in ECOTRIM

Univariate Methods
Without related series:
Boot, Feibes and Lisman - FD (first differences)
Boot, Feibes and Lisman - SD (second differences)
Denton Univariate
With related series:
AR (1) (maximum likelihood) – Chow Lin
AR (1) (minimum sum of squared residuals) – Chow Lin
Fernàndez
Litterman (maximum likelihood)
Litterman (minimum sum of squared residuals)

- ***Multivariate methods in ECOTRIM:***

Multivariate temporal disaggregation of time series with multivariate methods ensures the respect of both temporal and contemporaneous aggregation constraints.

Ecotrim proposes both adjustment and optimal techniques, in the least squares sense. Adjustment techniques require preliminary series fulfilling the temporal constraint. The accounting constraint is then achieved with respect to both temporal and contemporaneous aggregation constraints (in this case too, Ecotrim proposes either adjustment or optimal techniques, in the least squares sense).

Multivariate Methods
White Noise
Random Walk
Rossi
Denton Additive First Differences
Denton Additive Second Differences
Denton Proportional First Differences
Denton Proportional Second Differences

From the kick-off meeting and under the recommendation of Prof. Tommaso Di Fonzo, Benchmarking and temporal disaggregation methods expert, the set of methods will be redistributed in JECOTRIM. The table below shows the update.

Table 2: Methods in JECOTRIM

Univariate Methods
Benchmarking methods:
Modified DENTON
Regression-based temporal disaggregation:
Chow Lin
Fernàndez
Litterman
Reconciliation of systems of time series
Two-Step procedure (Quenneville & Rancourt 2005 – Di Fonzo & Marini 2009)
Balancing method
RAS with Plus/Minus option (negative values)

Why this choice?

The field of balancing/benchmarking time series has a long history with very large literature. The problem is usually the reconciliation of economic data to satisfy a set of constraints; it concerns Univariate and multivariate methods. Some pioneer papers in the domain are the ones published by Stone and al. (1942), Denton (1971), Chow and Lin (1971), Chollette(1984), etc. The most recent works were accomplished by Quenneville and Fortier (2006), Di Fonzo and Marini (1990, 2005, 2009).

The different reconciliation techniques use two popular approaches: a Fully-Matrix Regression-Based approach based on the least squares techniques, and a Recursive State-Space Regression Matrix approach. The first approach uses matrix calculation to provide exact results of parameters estimation while the second one uses iterative methods in which the estimation results depend on the starting points of the algorithm. Such points are usually generated using diffuse distribution or similar techniques that may introduce a bias on the final results. However, the two approaches should provide identical results if the starting points are exact ones. Choosing one or the approach is a major factor that may affect the quality of the final results.

Professor Tommaso DI FONZO how advised to use the Least Square approach for the implementation of the new ECOTRIM software. The reasons are double: first, the large progress in the capabilities of computing technology (sparse matrices, etc.) allow the manipulation of large system of data in very short time, and second, the least squares approach provides exact results

There is finally a lot of difference between the two libraries and a program of training will have to be prepared for a transition from the former set of methods to the new one.

Only the regression based methods are in common between ECOTRIM and JECOTRIM:

- DENTON in ECOTRIM was bugged and will not be comparable to the modified DENTON in JECOTRIM,

2-Step procedure and RAS with Plus/minus option are new. The 2-Step procedure is not comparable to the multivariate methods of ECOTRIM.

Annex 1: AGGREGATION ORDER

aggr	Description
2	Annual to half-yearly or half-yearly to quarterly
3	Quarterly to monthly
4	Annual to quarterly
6	Half-yearly to monthly
12	Annual to monthly

Annex 2: TYPE OF AGGREGATION

typeaggr	Description
flow	Sum
index	Average
last	Sample of the last
first	Sample of the first

Annex 3: SUMMARY TABLE OF STATISTICAL CONCEPTS

General

Number of valid cases	**Number of observations of temporally aggregated time series**
Degrees of Freedom	Number of observations (N) less the estimated coefficients (k)
Coefficient of determination (Buse, 1973)	It is a generalized R-squared statistic proposed by Buse (1973) in case of GLS estimation
Adjusted R-squared	Measure that imposes a small penalty in R-squared when a variable is added to the model
Standard Error of Regression	Measure of variability
Sum of Squared Totals	Measure of the total variation in dependent variable
Sum of Squared Residuals	Measure of the unexplained variation of the dependent variable
Sum of Squared Estimates	Measure of the explained variation of the dependent variable
Log-likelihood	Log-likelihood function of the model
F-statistic	Calculates F-statistic only if there are more than 1 regressor
Probability (F-statistic)	Displays the p-value corresponding to the reported F-statistic. Measures the significance of the F-statistic
Akaike Information Criterion	Measure of the explanatory capability of the model proposed by Akaike (1974)
Schwarz Information criterion	Measure of the explanatory capability of the model following a Bayesian approach suggested by Schwarz (1978)
Durbin-Watson statistic	Measure of the first-order autocorrelation

RELIABILITY INDICATORS

Calculated as follows:

$$\text{R.I} = \left({}^{\hat{\sigma}_{\hat{y}}}\!/_{\hat{y}_t} \right) \times 100$$

where $\hat{\sigma}_{\hat{y}}$denotes the estimated standard error for the estimated series and $\hat{y}_t$the estimated series. Notice that JEcotrim does not performs the percentage.

INTERVAL ESTIMATION

The lower limit value of the confidence interval is calculated as follows:

$$\hat{y}_t - t_{0,95}^{n-k} \cdot \hat{\sigma}_{\hat{y}}$$

being $t_{0,95}^{n-k}$ the value of the t-Student table, for the (n-k) degrees of freedom. Whilst the upper limit as

$$\hat{y}_t + t_{0,95}^{n-k} \cdot \hat{\sigma}_{\hat{y}}$$

GROWTH RATES

The growth rates performed by JEcotrim are two:

– The lag 1 growth rate: T1G= $\left(\frac{\hat{y}_t}{\hat{y}_{t-1}} - 1\right) \times 100$

The annual growth rate: TAG= $\left(\frac{\hat{y}_t}{\hat{y}_{t-s}} - 1\right) \times 100$ where s=4. Again, the JEcotrim interface does not perform the percentage.

Annex 4: BENCHMARKING ESTIMATION METHODS

bench	Description
AFD	Additive First Differences
ASD	Additive Second Differences
PFD	Proportional First Differences
PSD	Proportional Second Differences

Annex 5: SUMMARY TABLE OF DESCRIPTIVE STATISTICS

Number of variables	Number of variables
Frequency conversion	Frequency conversion
Number of LF periods	Number of LF periods
Number of HF periods	Number of HF periods
Number of contemporaneous constraints	Number of contemporaneous constraints
Dimension of the linear system	Dimension of the linear system
Density of the system matrix (%)	Density of the system matrix (%)
Number of constraints	Number of constraints
Method of reconciliation	Method of reconciliation
Precision level	Precision level
Absolute average discrepancy	Absolute average discrepancy