eurostat

Doc. Eurostat/A4/SA/02/doc
Available in **EN**
04/02

Informal working group on Seasonal Adjustment

Fifth meeting

Luxembourg, 25 – 26 April 2002 at 9 h 30

Room Quetelet, Bech building

Item 3.3

New capabilities of Demetra v2.0 SP1

New capabilities of Demetra v2.0 SP1

Jens Dossé, OECD[1], and Jean-Marc Museux, Eurostat March/April 2002

Introduction

This document describes the most important improvements and new experimental functions in Demetra version 2.0 Service Pack 1, the seasonal adjustment interface to Tramo/Seats and X-12-Arima, developed by the Jens Dossé (formerly in Planistat Group) for Eurostat.

Former versions of Demetra were limited to lonely execute the interfaced seasonal adjustment methods with the parameters provided by the user or proposed by Demetra. These options might not have been the optimal ones for each time series (e.g. the number of trading day regression variables), and for rejected adjustments, the user had to change manually some of the options to obtain better modelling results. Comparisons between two results of one method were time consuming because several diagnostic statistics had to be weighted manually. The model obtained from automatic modelling with Tramo/Seats or X-12-Arima largely depended on the series span used. Thus, the new Demetra version besides basic enhancements proposes experimental features for improving model quality and stability and increasing user-friendliness for easier modelling even for large-scale sets of series. These functions can be activated by the users by switches in the graphical interface. If switched off, these new features do not affect the regular run of Demetra

A adjustable composite index combining existing quality measures

Demetra provides tables with the necessary statistics on modelling and decomposition allowing a quality control of the adjustments. One of the tables is called “Information on Diagnostics”. It contains standard statistics in the case of an “Automated Module” project, and a configurable degree of detail for “Detailed Analysis” projects.

Information on Diagnostics	Model 1 (Tramo-Seats)	Model 2 (Tramo-Seats)
STATISTICS ON RESIDUALS		
Ljung-Box on residuals	23.88 [0, 33.90] 5%	23.33 [0, 33.90] 5%
Box-Pierce on residuals	0.33 [0, 5.99] 5%	0.26 [0, 5.99] 5%
Ljung-Box on squared residuals	25.02 [0, 33.90] 5%	22.99 [0, 33.90] 5%
Box-Pierce on squared residuals	0.98 [0, 5.99] 5%	0.94 [0, 5.99] 5%
DESCRIPTION OF RESIDUALS		
Normality	1.89 [0, 5.99] 5%	2.04 [0, 5.99] 5%
Skewness	0.03 [-0.42, 0.42] 5%	0.05 [-0.42, 0.42] 5%
Kurtosis	3.59 [2.16, 3.84] 5%	3.61 [2.16, 3.84] 5%
OUTLIERS		
Percentage of outliers	0.00% [0%, 5.0%] ad-hoc	0.00% [0%, 5.0%] ad-hoc

Figure 1

Figure 1 shows such a table as produced by previous Demetra versions. There, to compare the quality of two different models, the user had to compare each diagnostic statistic and according to his own practice, weight between them. The new “SA quality index” simplifies and speeds up such tasks, and also helps to automate the model comparisons needed in the other new procedures as described later.

How is this quality indicator built?

[1] The views expressed in this document do not necessarily reflect the official views of the OECD.
Address: OECD; 2, rue André Pascal; 75775 Paris cédex 16; France; E-mail: jens.dosse@oecd.org

Generally spoken, the “SA quality index” is constructed as a sum of all diagnostic statistics that are also applied in the usual Demetra quality check.

$$\sum_j statistic_j$$

Since these statistics have different sizes in terms of optimal and critical values, they needed to be normalised.

$$\sum_j \left| \frac{statistic_j - optimalvalue_j}{criticalvalue_j - optimalvalue_j} \right|$$

The critical values are chosen in accordance with the Demetra quality check. If a diagnostic statistic exceeds the critical value, the Demetra quality check for this series will also fail and the adjustment will be rejected. Furthermore, to allow for more flexibility, the normalised statistics can be exponentiated (penalty option) and weighted in the sum, but the default values for these options are 1, which have no effects.

$$\sum_j weight_j \cdot \left| \frac{statistic_j - optimalvalue_j}{criticalvalue_j - optimalvalue_j} \right|^{penalty}$$

Because different numbers of diagnostics statistics might be used in the quality check, the overall statistic is normalised again to a “critical value” of 10. Demetra’s fully customisable formula is thus:

$$QualityIndex = \frac{10}{M} \cdot \sum_j weight_j \cdot \left| \frac{statistic_j - optimalvalue_j}{criticalvalue_j - optimalvalue_j} \right|^{penalty}$$

M is the numbers of diagnostics statistics used in the quality check. The weights for each diagnostic statistic as well as the penalty option can be modified in a corresponding Demetra dialog.

Demetra’s default formula is:

$$QualityIndex = \frac{10}{M} \cdot \sum_j \left| \frac{statistic_j - optimalvalue_j}{criticalvalue_j - optimalvalue_j} \right|$$

The values of the “SA quality index” can vary from 0 (optimal value) to infinity (worst value). Each normalised diagnostic has a “critical value” of 1, thus if the “SA quality index” exceeds the “critical value” of 10, at least one statistic must have been significant, and thus the adjustment rejected. The “SA quality index” of automatically accepted adjustments is thus always smaller than 10. It is now shown in the first row of the table of “Information on Diagnostics” (see figure 2).

Information on Diagnostics	Model 1 (Tramo-Seats)
SA quality index (stand. to 10)	2.122 [0, 10] ad-hoc
STATISTICS ON RESIDUALS	
Ljung-Box on residuals	23.88 [0, 33.90] 5%
Box-Pierce on residuals	0.33 [0, 5.99] 5%
Ljung-Box on squared residuals	25.02 [0, 33.90] 5%
Box-Pierce on squared residuals	0.98 [0, 5.99] 5%
DESCRIPTION OF RESIDUALS	
Normality	1.89 [0, 5.99] 5%
Skewness	0.03 [-0.42, 0.42] 5%
Kurtosis	3.59 [2.16, 3.84] 5%
OUTLIERS	
Percentage of outliers	0.00% [0%, 5.0%] ad-hoc

Figure 2

The “SA quality index” is also reported in the log file for each shortened model in the Stability analysis if the corresponding printing option is set.

Remark: For most Tramo/Seats residual diagnostic statistics the first test critical value in the Demetra quality check is that for a 95% confidence level (default). That critical value is not critical in the sense of the fail of the Demetra quality check. In order to reject an adjustment because of one single significant residual diagnostic statistic, the value must exceed the critical value that corresponds to a 99.9 % confidence level (default). Thus the "SA quality index" is always based on critical values that correspond to a 99.9% confidence level (default).

Optimisations (reduction) of the number of trading-day regressors

Neither Tramo/Seats nor X-12-Arima are able to decide for the most appropriate number of trading day regressors for a time series. Practical experience has shown that models (e.g. with unsatisfactory results) can often be improved by changing the number of trading day regressors included in the regARIMA modelling.

Thus, the idea of the new Demetra function is to simplify the task of a user to test for the performance with the different trading day options of the seasonal adjustment methods. When the "Optimisation of the number of trading day regressors" is switched on, then for all currently selected time series, Demetra automatically runs adjustments with stepwise reduced numbers of trading day regressors, compares their diagnostic statistics (in terms of the "SA quality index"), and chooses for the best option.

Remark: This function is only active, if the pre-test for the trading day effect is also selected.

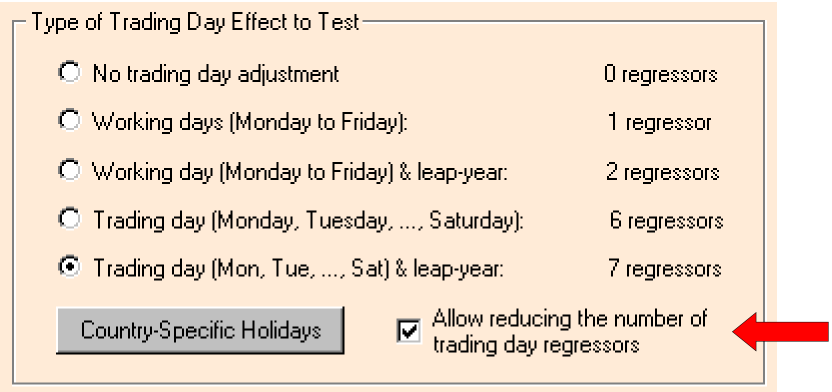

Figure 3

Expert system for rejected adjustments

Since its very first version, Demetra emphasised on the importance of quality of model used in the seasonal adjustment procedure, and provided checks for the acceptance of the result of a modelling. Depending on the type of data, a relatively important number of series can show difficulties concerning their modelling. Figure 4 shows the hints that were (and are still) given, in the case the adjustment was rejected, for manual parameter modifications that might help the user in finding better performing models. Those hints were proposed by the Seasonal Adjustment team of Eurostat, and already elaborated during the specification phase in the Demetra development.

It is very clear that the manual treatment of single series with insufficient modelling especially for large-scale datasets is a very painful and time-consuming task for the user. Considering that Demetra was intended to improve the general seasonal adjustment quality, but also being user-friendly and automate processing to increase productivity, it was desirable that such repeating and lengthy tasks were available as an automatic function.

That functionality of automatically trying to improve the modelling of a series with a rejected adjustment is finally added in the service pack 1 of Demetra version 2.0. It is called "Expert System" and "mimics" the usual actions of a user trying to achieve the improvements. It actually implements the Demetra's hints as shown in figure 4.

For all newly rejected adjustments during an adjustment run the “Expert System” analyses the series structure, the diagnostic statistics from that run and the automatic modelling parameters used. Based on that information, it fixes the otherwise automatic parameters to one of the possible options and executes the adjustment and quality check in a loop until a convenient model has been found:

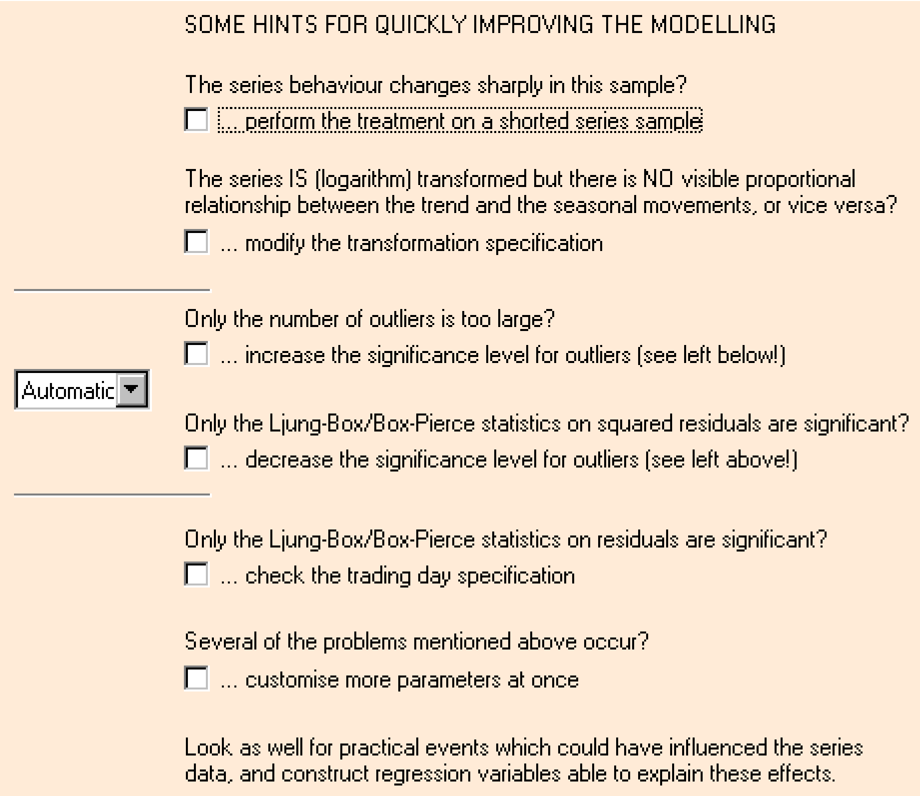

Figure 4

The “Expert System” refines the following automatic parameter options:

- Pre-test for the log-transformation (multiplicative modelling): For the (currently) linearised series (which is outlier corrected) as well as the log-transformed linearised series, Demetra calculates a simple moving average and a moving range (mean distance from moving average) with lengths of one year, for which it receives the correlation coefficients and the test statistic for linear relationships. If the test statistic for the linearised series is higher than the one for the log-transformed linearised series then the multiplicative model is used and vice versa. If the linearised series is not available than the original series is used instead. After the test, the log-transformation option is set accordingly to the results of this test and the SA methods are re-executed if necessary.
- Pre-test for the correction of trading days and the optimisation (reduction) of the number of trading day regressors was allowed: Demetra compares the diagnostic statistics for the adjustments with stepwise-reduced numbers of regressors and use the trading day option that corresponds to the best “SA quality index”.
- Automatic critical value for the detection/correction of outliers: Demetra tries to optimise this value depending on the number of outliers currently found as follow:
  - If there are too many outliers, then the critical value for outliers is increased up to the smallest t-value of all outliers and the methods are run again.
  - If there are not too many outliers then the critical value is decreased to the smallest possible value (Tramo: 2.0, X-12-Arima: 2.8). This step is only done once.
- Automatic identification/selection of the ARIMA model: Demetra loops through a certain set of combinations of ARIMA model orders: The regular autoregressive and moving average orders p and q are varied from 0 to 3 with the condition that $p+q<4$. All the other model orders are held fixed at the values of the first regular run. For each ARIMA model type, the procedure for improving the VA value is called.

- If nothing else helped and if specified so in the Expert System options, Demetra will stepwise (year-wise) cut the modelling span at the beginning of the time series down to minimal 4 years, and stop if an acceptable adjustment is found.

The implementation of the “Expert System” is completely similar between Tramo/Seats and X-12-Arima with one exception: The algorithm for testing different combinations of regular ARMA orders starts for X-12-Arima with a special Tramo run with an automatic ARIMA model identification. The “Expert System” does not replace any algorithms implemented in the SA methods. Since the mean correction parameter does not provide an automatic option (neither in Tramo nor in X-12-Arima), it is not modified by the “Expert System”. The first experiences have shown that it would be the best that the user performs a full run of the “Expert System” without mean correction after a previous unsuccessful run of the expert system with mean correction (that is especially the case for X-12-Arima since it does never change this parameter itself).

For very difficult series, the procedure can be a lengthy process (up to 1 minute per series), but the user is freed from that work. See the tables in the Benchmarking title below for more information on the success rate of the “Expert System”.

(Model) Stability analysis

As a study form Estelle Renaud (Planistat) for Eurostat on the stability of the regARIMA modelling showed, the model obtained from automatic modelling with Tramo/Seats largely depends on the series span used. For that reason Demetra proposes a Model “Stability Analysis” that
- Runs the seasonal adjustment method with the same automatic parameter options but with a stepwise (per period) shortened modelling-span (up to one year).
- For each run finds the newly identified modelling parameters (ARIMA model orders, log-transformation, mean correction, trading day and Easter effect correction, X11 seasonal and trend filter) and calculates the “SA quality index”, see figure 5.
- Fixes the modelling span to the *longest* time interval that corresponds to the model with *acceptable* results, *best* “SA quality index” and with the modelling parameters *most often* identified.

The model choice in the “Stability Analysis” relays thus on an algorithm that is based on the “SA quality index” and additional weighting terms punishing too “old” models (with shorter modelling spans) and giving advantage to models those parameters were more often identified. That is done in the following way:

Punishment of shorter modelling spans:

0 periods cut	à historicweight = $1+(k \cdot mq)$
1 period cut	à historicweight = $2+(k \cdot mq)$
…	
(*mq*-1) periods cut	à historicweight = $mq+(k \cdot mq)$

The sum of all historicweight’s: $(mq+1)\cdot\frac{mq}{2}+k\cdot mq^2$

When a “new” acceptable model (unique ARIMA model orders, mult./add.type, mean correction, trading day correction, Easter correction, and X11 seasonal and trend filters) is found during the run, it is attributed a maximum value of

$$10\cdot(mq+1)\cdot\frac{mq}{2}+k\cdot mq^2$$

For each similar (acceptable) model that is found in the run, the term

$$(10-QualityIndex)\cdot((mq-CutPeriods)+k\cdot mq)$$

is subtracted.

The parameter *k* specifies the importance of history. Its default value is 0 that means that no additional importance is granted to historical models.

The final value obtained for each type of model is used in the choice for the best model in the "Stability Analysis". For the best model type the model with the longest modelling span is used.
Figure 5 shows the text output produced by the "Stability Analysis".

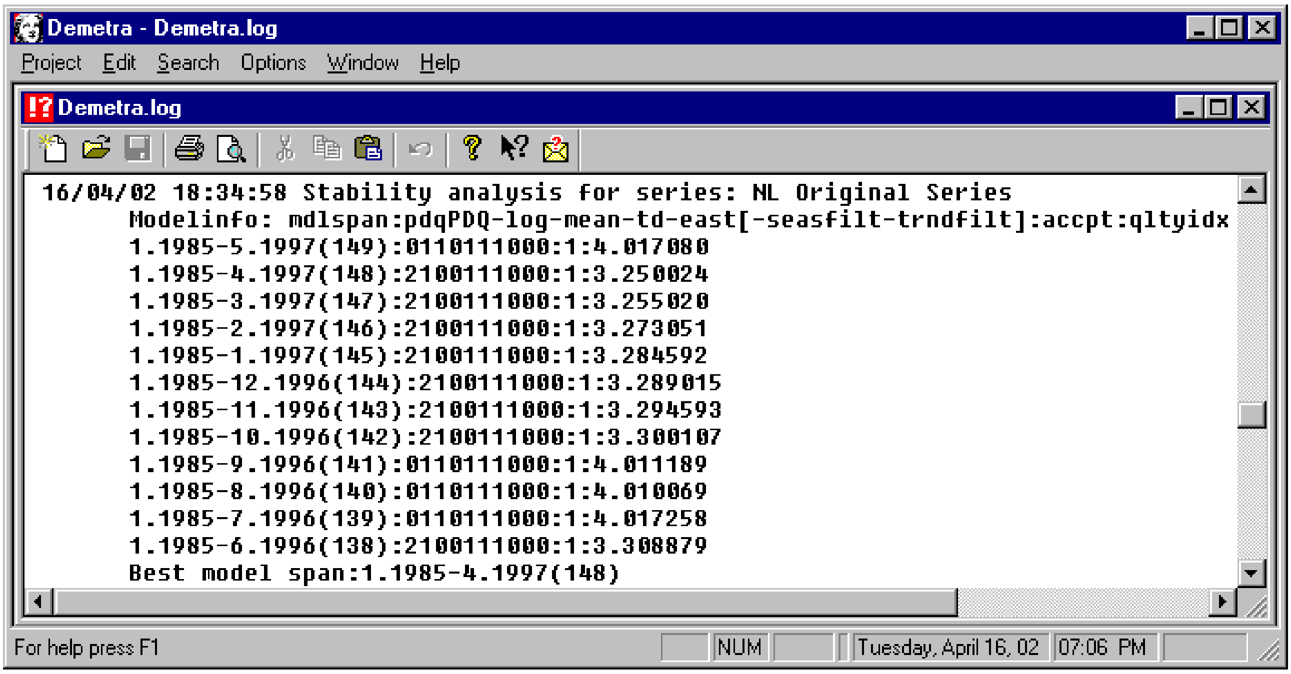

Figure 5

Benchmarking , and

An empirical analysis of the performance of the 3 new tools Optimisations (reduction) of the number of trading-day regressors, Expert system, and (Model) Stability analysis returned the following results:

- The number of rejected series could significantly be reduced by using these 3 tools:

Periodicity	Method	Total n° series	Number of rejected series				
			Normal adjustment		+	+ +	+ + (no mean)
Monthly	Tramo/Seats	11786	3079	1696	1418	948	-
	X-12-Arima	11786	6311	4636	3871	2788	2504
Quarterly	Tramo/Seats	687	98	24	19	2	-
	X-12-Arima	687	247	116	108	82	68

Table 1

- The use of the new tools increases the average adjustment times:

Periodicity	Method	Average adjustment times (in seconds per series)				
		Normal adjustment		+	+ +	+ + (no mean)
Monthly	Tramo/Seats	1.4	2.3	6.2	32.9	-
	X-12-Arima	1.4	7.5	24.8	62.6	62.2
Quarterly	Tramo/Seats	0.1	0.6	1.2	3.3	-
	X-12-Arima	0.7	3.0	4.8	14.0	14.2

Table 2

Because this information largely depend on the PC configuration (e.g. processor speed), the table is here repeated using figures expressing the increase of average adjustment times:

Periodicity	Method	Increase of average adjustment times (ratios)				
		Normal adjustment		+	+ +	+ + (no mean)
Monthly	Tramo/Seats	1.0	1.7	4.5	23.6	-
	X-12-Arima	1.0	5.4	17.8	45.0	44.7
Quarterly	Tramo/Seats	1.0	6.3	11.9	32.2	-
	X-12-Arima	1.0	4.5	7.3	21.2	21.4

Table 3

- The Trading Day Optimisation (and the Stability Analysis) can improve the seasonal adjustment quality even for already accepted series:

Periodicity	Method	N° of accepted series	SA quality index (mean)			SA quality index (variance)		
			Normal adjustment		+	Normal adjustment		+
Monthly	Tramo/Seats	8717	2.6	2.1	2.2	0.6	0.4	0.4
	X-12-Arima	5475	3.9	3.5	3.6	1.6	1.2	1.3
Quarterly	Tramo/Seats	589	1.9	1.4	1.4	0.5	0.3	0.3
	X-12-Arima	440	3.2	2.7	2.7	1.2	0.9	0.9

Table 4

- The Expert System also improved the seasonal adjustment quality for still rejected series:

Periodicity	Method	N° of rejected series	SA quality index (mean)		SA quality index (variance)	
				+		+
Monthly	Tramo/Seats	972	20.8	7.2	9463.3	97.1
	X-12-Arima	2504 (2315)	17.9	11.7	12996.8	4967.9
Quarterly	Tramo/Seats	2	20.8	6.1	262.2	0.0
	X-12-Arima	68	7.9	5.8	39.4	5.6

Table 5

For this analysis an original data set from Eurostat including business series from different areas (production index, turnover, employment, wages, and consumer prices). The monthly price series showed often very special behaviour especially invariant data during several months that explains the relatively high percentages of rejected adjustments in all steps of the analysis.

Remark: Since the calculation of the SA quality index for the SA methods Tramo/Seats and X-12-Arima are not identical, the statistics are not comparable between each other.

Business-cycle analysis and long-term trends

In the working document nº 0108 of the Banco de España "Time Aggregation and the Hodrick-Prescott Filter", Agustín Maravall and Ana del Río describe how the seasonal adjustment method Seats can be used to produce Hodrick-Prescott-filter equivalent business-cycles by means of an ARIMA (0 2 2)(0 0 0) model-based decomposition.

Demetra 2.0 SP1 contains an implementation of this methodology that can be freely used to automatically calculate business-cycles and long-term trends from any series.

The user can specify the Hodrick-Prescott filter parameter λ (lambda) for one of the 3 time series periodicities (see the 3 edit fields in figure 6), or use the default values.

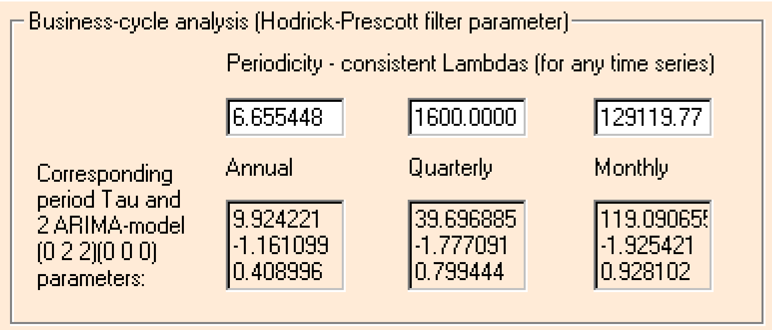

Figure 6

The above named document of Agustín Maravall and Ana del Río shows that the parameter λ (lambda) biuniquely corresponds to a maximum length (in number of periods) of the cycles filtered. E.g. for a quarterly time series, a value λ=1600 gets a filter that produces a business-cycle component that contains cycles with a length of up to approximately 40 periods that corresponds to approximately 10 years. The maximum cycle length (in number of periods) is called τ (tau).
If we annualise that original series (annual time-aggregate) and want to obtain the equivalent annualised business-cycle component through the HP filter, we must also convert the maximum cycle length τ (tau) to 10 periods (still representing 10 years). The filter parameter λ (lambda) that corresponds to that τ is approximately 6.7.

In that way, Demetra obtains periodicity-consistent λ's (lambda's) for the needed time series periodicity based on the given parameter λ (lambda) (even if that time series periodicity is not included in the 3 examples for annual, quarterly or monthly series).
Afterwards, Demetra calculates the 2 moving average coefficients for the ARIMA-model (0 2 2)(0 0 0) that correspond to the Hodrick-Prescott filter parameter λ (lambda), as described in the working document mentioned above.

The ***Business-Cycle*** series is obtained by sending the *Forecasted Stochastic Trend-Cycle* series from a first normal Tramo/Seats run together with the ARIMA (0 2 2)(0 0 0) model derived from λ (lambda) to Seats and recovering the *Irregular Component*.

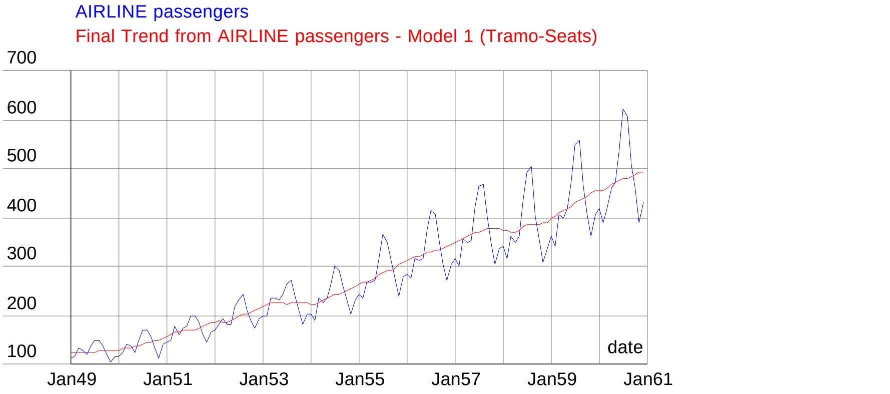

Figure 7

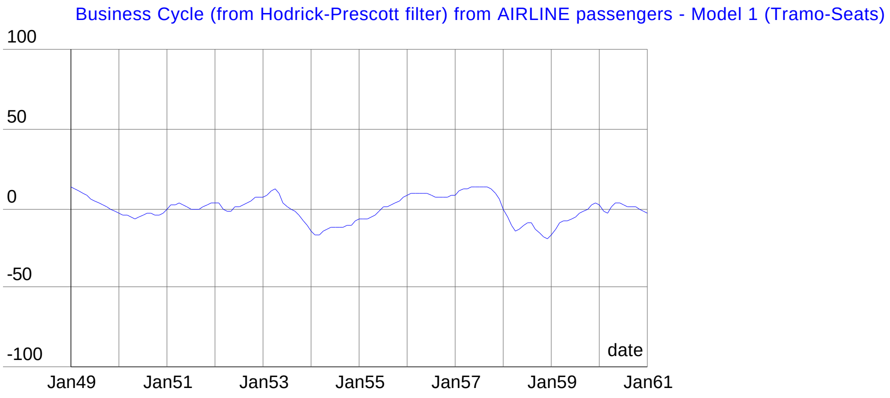

Figure 8

The ***Long-Term Trend*** is the difference between the *Final Trend-Cycle* from the first run and the *Business-Cycle*.

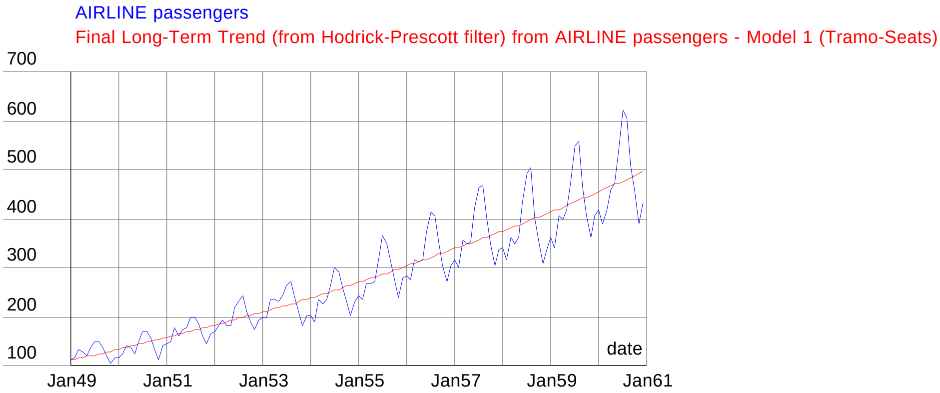

Figure 9

The "***Business-Cycle + Irregular***" series is the difference between the *Final SA* series from the first run and the *Long-Term Trend*.

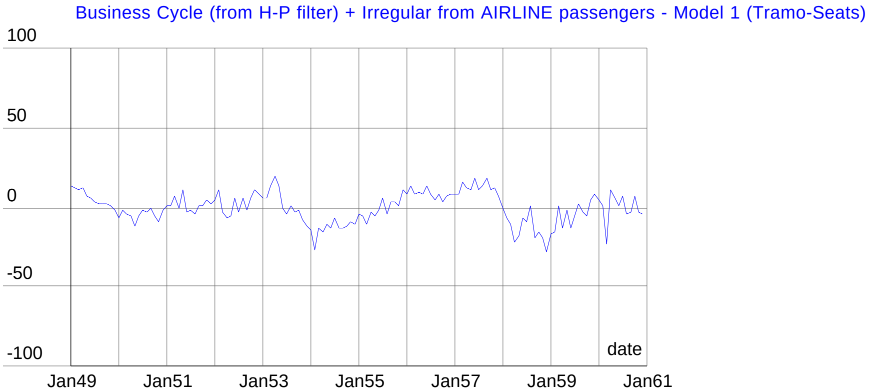

Figure 10

Remarks:

- Level-shifts are contained in the Long-Term Trend, and thus not in the "Business-Cycle + Irregular" series.
- The default value in Demetra for the parameter λ (Lambda) for quarterly series is 1600 that is a "de facto industry standard" (European Central Bank (2000)).

Basic enhancements

Demetra 2.0 SP1 offers new and often requested features improving data accessibility, productivity, flexibility, and user-friendliness:

- A direct access to data in Oracle Express databases is provided. The supported software includes Personal Express and Express Server 5 and 6.
- The access to FAME databases has been improved. The speed of the data catalogue retrieval has been significantly increased. The security in the Demetra Fame server (Unix version) has been improved.
- Parameters can be saved in most of the interfaced databases using the original formats of Tramo/Seats or X-12-Arima.
- The modelling with Tramo/Seats can be limited to a user-defined time span (like the modelspan argument in X-12-Arima).
- Graphs with the seasonal factors/component per period can be plotted (see figure 12).

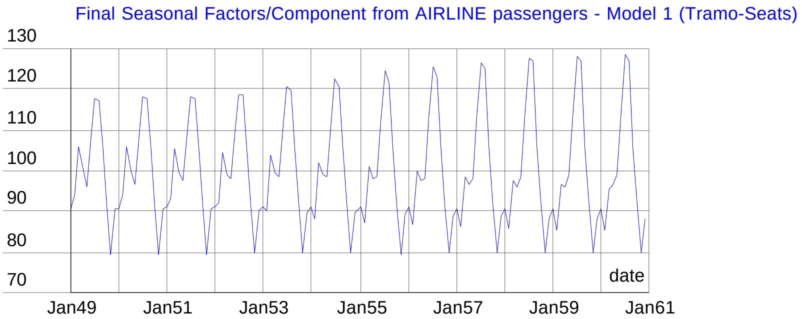

Figure 11

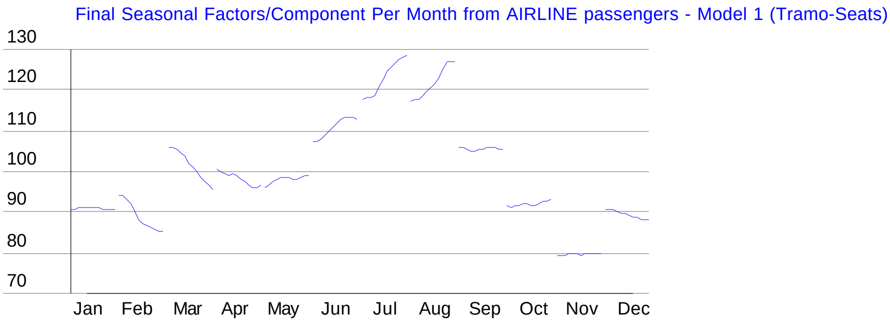

Figure 12

- Models in the Detailed Analysis Module can easily be deleted/executed/cut/copied/pasted.

Conclusion

For many series, the previously available automatic tools in Demetra were not sufficient to assure a fully automated modelling. Considering that Demetra is a tool for statistical production and thus for the treatment of large-scale set of time series, the necessity of too much manual user interactions for model improvements was inadmissible. Thus, new experimental tools for automatic optimisation of the trading day correction, the model span and for the improvement of rejected adjustments were implemented into Demetra that should further evaluated but should facilitate the work in seasonal adjustment productions. Most of the underlying ideas of the algorithms were already contained as "Hints" in the corresponding screen of Demetra since its very first version. Also the practical evaluations of the new tools tend to prove their usefulness not only in respect to automating the improvement of the modelling of series with rejected adjustments, but additionally in respect to improving the quality of previously accepted adjustments. The tools significantly increase the processing times but those are negligible in comparison to manual treatments that were necessary in former versions of Demetra.

The possible next steps of the Demetra development could possibly include the following improvements:

- Implementation of new scientifically well founded, harmonised statistical tests to measure the different aspects of a high-quality seasonal adjustment and its incorporation into the "SA quality index"
- Improvement of the holiday tools (to correct trading day regressors): use in X-12-Arima, movable feasts for a wider field of religions, direct input of holiday variables.
- Common analysis tools (aggregation analysis, sliding spans analysis, ...), especially the user-friendly presentation of their results in the Detailed Analysis Module of Demetra. Hereby must be considered, the rapid development of the US Bureau of Census package in collaboration with Agustin Maravall. Demetra itself must concentrate on its real added value:
  - Efficient, automatic and high-quality modelling and adjustment of large-scale sets of series from different sources
  - User-friendly summaries and the presentation of the results of all analysis parts