SSE-DP-2024-4

On SarSIML

(A Seasonal Adjustment Method)

Seisho Sato

(University of Tokyo)

and

Naoto Kunitomo

(Institute of Statistical Mathematics)

November 2024

On SarSIML

(A Seasonal Adjustment Method) *

Seisho Sato †
and
Naoto Kunitomo ‡

November 14, 2024

Abstract

We explain a new seasonal adjustment program called SarSIML (or S-SIML). It is based on the (real-valued) spectral decomposition of non-stationary time series, which is an application of the SIML filtering method developed by Kunitomo and Sato (*The SIML Filtering Method for Noisy Non-stationary Economic Time Series*, 2024, JSS-Springer Series, Springer, forthcoming).

Key Words

Seasonal Adjustment, SarSIML (S-SIML), Frequency Domain Decomposition, Noisy non-stationary time series, Macro-economic time series,

*2024-11-14. A preliminary memorandum.

†Graduate School of Economics, University of Tokyo, Bunkyo-ku, Hongo 7-3-1, Tokyo 113-0033, seisho@e.u-tokyo.ac.jp

‡The Institute of Statistical Mathematics, Midoricho 10-3, Tachikawa-shi, T okyo 190-8562, JAPAN, kunitomo@ism.ac.jp

1 Introduction

Over time, a large volume of economic time series data is observed. Especially in major macroeconomic official data, due to reasons related to data collection and creation, measurements of data are taken at various frequencies and timings, such as daily, monthly, quarterly, and annually, and these data are published by different government departments in the ministries of Japan. Each series, including key macro time series such as consumption, investment, government spending, and imports/exports, has its unique characteristics, which is measured by nation-wide statistical survey. They are not always published in a mutually adjusted manner.

For purposes such as understanding overall economic trends, conducting data analysis, and making policy decisions, these conditions surrounding data creation and publication are not ideal. For macroeconomic series that indicate the overall economy in a nation, it is preferable to produce and publish accurate data as promptly as possible to understand the latest macroeconomic conditions. To create nationwide data, however, significant amount of sample-survey data needs to be collected. For example, in Japan, GDP and its main components are published quarterly, with time lags we need to incorporate various data, undergoing revisions in preliminary and secondary reports, and we have final updates. Data on the investment series becomes available after the preliminary GDP report, using the underlying data collected in a large scale, which believed to be more accurate than the data used for the preliminary estimate of GDP, for instance. Then, it leads to revisions of GDP in the initially published quarterly series. Errors arising from such data revisions should be treated as observation errors in the estimation process of true state variables. Inconsistencies sometimes arise, such as between published quarterly data and the latest monthly series obtained from sample survey data. Additionally, in Japanese official statistics, each department within the central government is responsible for different series, adding complexity due to separate responsibilities in publication. It seems that in Japan the Statistics Bureau does not have much power to overrule the individual practice in different ministries.

In many countries, including Japan, seasonal adjustment methods such as X-12-ARIMA and X-13ARIMA-SEATS, based on moving averages, are widely used. These seasonal adjustment programs use a univariate time series regression model called the Reg-ARIMA model for internal calculations. Practically in Japan, each department in ministries individually applies and operates the Reg-ARIMA model on each series they publish. Therefore, despite various relationships existing among macro time series, consistent seasonal adjustment across multiple time series is not always ensured as time series component decomposition is necessary due to the mixture of different structural elements.

This memorandum introduces a seasonal adjustment method based on the SIML filtering approach from a statistical time series analysis perspective, without adhering to these existing conditions. It explains the usage of the “sarsiml” method (also known as S-SIML seasonal adjustment) using R programming. The SIML filtering method, developed by Kunitomo and Sato (2021, 2024), allows for state estimation of trends, cyclical components, seasonal components, irregular components, and even taking account of breakpoints in noisy non-stationary time series. For Japanese macro-consumption series, Kunitomo, Sakurai, and Sato (2022) provided a detailed explanation of the SIML filtering method. Using this SIML method, direct estimates of trends, cyclical components, seasonal components, and seasonally adjusted series can be obtained from non-stationary and seasonal quarterly or monthly series through frequency decomposition. We suggest a set of parameters in the SIML filtering method so that the SarSIML program can be applied to real data in an automatic way.

2 Explanation of How to Use the R Program SarSIML (S-SIML) for Seasonal Adjustment

This program, executable on the free software R, is a simplified version of the x12siml R program originally developed by Seisho Sato. The x12siml program provides a broader range of functions for analyzing economic time series data with seasonality and non-stationarity. (Refer to Sato (2023) for more details [1].)
(i) Place the file in the directory where R is launched, and in R, execute:
> source("sarsiml_ver1.1.R")
This completes the installation. (Alternatively, you can load the R source file directly in text format.)
(ii) Next, to execute with quarterly data, use:
> sarsiml(quarterlydataname)
or, for monthly data:
> sarsiml(monthlydataname, frequency=12)
(iii) The data file can be loaded using R commands. Alternatively, if using an Excel file, data can be loaded directly by copying from Excel and running:
> data = read.delim("clipboard")
(iv) The trend parameter m and seasonal fluctuation parameter $sorder$ are set automatically ($sorder$ =1 and, m on a two-year cycle for monthly data and three years cycle for quaterly data). However, they can be manually configured as necessary by consulting resources like Kunitomo, Sakurai, and Sato (2022) or the x12siml manual. When $sorder$=1, the model excludes three frequencies from the bandwidth range around the seasonal cycle λ_s (excluding $\lambda_s = 1/2$), that is $[\lambda_s - 1/(2n), \lambda_s + 1/(2n)]$, where n is the number of observations. For $\lambda_s = 1/2$, we use the frequency band $[1/2 - 1/n, 1/2]$. For cases when seasonality clearly changes, specific adjustments can also be made. For example,
> sarsiml(monthlydataname, frequency=12, trend=10, sorder=2)

[1] https://github.com/sato-labo/x12siml

When we use the default setting, sometimes seasonal fluctuation remains. One way to cope with such case, set
> smooth =T
and then components around the frequency $\lambda_s = 1/2$ are removed. It is possible to $sorder = 0$ when there is no seasonality.
(v) Results from the estimation can be extracted and reused as follows:
>res = sarsiml(shouhi)
>res$trend
>res$adj
>res$$Z$
>res$seasonal
>res$noise
(vi) The program is set by default as $pa = 2$, $pb = 2$ to ensure stable seasonal estimates near the beginning and end of the time series data. This corresponds to creating a forecast and back-cast series as in X-12-ARIMA using the ARIMA model. (We assume that the seasonal fluctuation follows a random walk model and the trend-cycle and seasonal components in the previous year and next year follow as in the first year and last year, respectively.)
By setting $pa = pb = 0$, you can disable the use of prior information, though stable estimates around the start and end of the series may not be obtained in many cases.
(vii) Two AIC values shown in the figure are pseudo-AICs derived from the Gaussian likelihood function based on the estimated noise, which are calculated from n observations and m observations after the effects of seasonality are removed. They are calculated from $\hat{\sigma}_z^2 = (1/n)\sum_{i\in \mathrm{I}} z_i^2$, where I is the index set of residuals, which corresponds to Akaike's Final Prediction Errors (FPE) as (number of observations) $\times$ $\log[\hat{\sigma}_z^2]$. They indicate the model's fit to the filtered data after removing seasonality. Specifically, the AIC calculated here serves as a measure of data fit, based on variance estimates from the transformed data matrix Z from all data or the data with seasonal components removed.

3 Examples

As an example, Figure 1 presents the analysis results of the quarterly real final consumption series (1994-2000, Cabinet Office). The solid line in the upper left graph represents the original series, with the red line showing the trend-cyclical component. The upper right graph shows the seasonal component estimated from the time series, while the central left graph shows the estimated noise component. Unlike X-12-ARIMA or X-13ARIMA-SEATS by the U.S. Census Bureau, this method estimates seasonality directly from frequency components, eliminating the need for Reg-ARIMA models. The lower left graph shows the Z series obtained through K_n-transformation, with seasonality appearing only around frequencies λ_s=0.25, 0.5 for quarterly macro time series with seasonality.

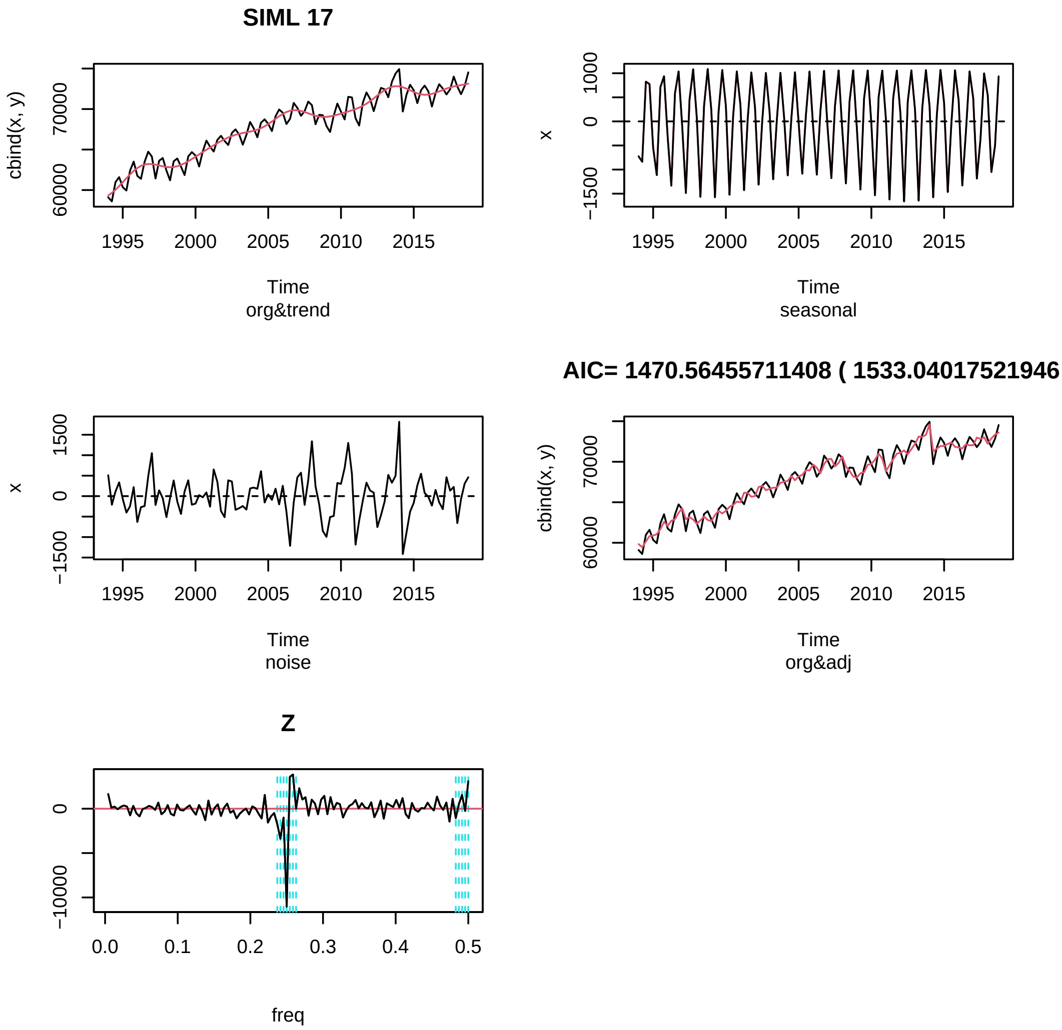

Figure 1： S-SIML Seasonal Adjustment (Real Consumption)

For an example with monthly data, Figure 2 illustrates seasonally adjusted values from household consumption data (2002-2019), as discussed in Kunitomo, Sakurai, and Sato (2022). Similar to the quarterly data, this figure includes the original series, seasonal component, irregular component, and seasonally adjusted values. The orthogonal component Z obtained through K_n-transformation exhibits effects around the frequencies corresponding to seasonal components. The lower left graph shows the Z series obtained through K_n- transformation, with seasonality appearing only around frequencies $\lambda_s = 1/12, 2/12, 3/12, 4/12, 5/12, 6/12$ for monthly macro time series with seasonality. Theses features may explain some of the confusion surrounding discussions on seasonality of monthly economic time series in the past.

4 Some Remarks

In this memorandum, we explained a method for obtaining trends, cyclical components, seasonal components, and seasonally adjusted series directly from non-stationary quarterly and monthly time series through frequency decomposition, which includes noise and seasonality. Detailed explanations are left to Kunitomo-Sato (2021), Kunitomo (2024), and Kunitomo, Sakurai, and Sato (2022). Frequency-based seasonal adjustment programs are rare; however, considering the goal of “removing seasonality,” frequency decomposition is a natural approach. Nevertheless, since most economic time series evidently include noisy non-stationary components, spectral decomposition with this in mind has not gained much attention.

As background for the method discussed here, it is worth mentioning on the effort of analyzing macro-consumption data in China. In Chinese official statistics, for some key macroeconomic indicators, it has been a long-standing practice not to publish January figures individually. Instead, combined figures for January and February are released later. This practice, possibly aimed at avoiding misunderstandings by accounting for seasonal variations such as the Spring Festival, is distinct from other major developed countries. The issue was pointed out by an economist in the Cabinet Office of Japan in our project of statistical consulting service at ISM (Institute of Statistical Mathematics).

In the United States, Japan, and many European countries, seasonally adjusted series are created to understand macroeconomic conditions, and they are released by media outlets. Programs like X-12-ARIMA and X13ARIMA-SEATS, developed by the U.S. Census Bureau, are widely used for this purpose, implicitly assuming the Gregorian calendar with an annual seasonality definition. However, in China, the influence of the lunar calendar is substantial, raising fundamental questions such as ”What is seasonality?” and ”What is seasonal adjustment?”

The S-SIML seasonal method described in this memorandum, as illustrated by Kunitomo (2024), can be applied to investigate noisy non-stationary seasonal economic time series even when missing values are present. Beyond the missing value

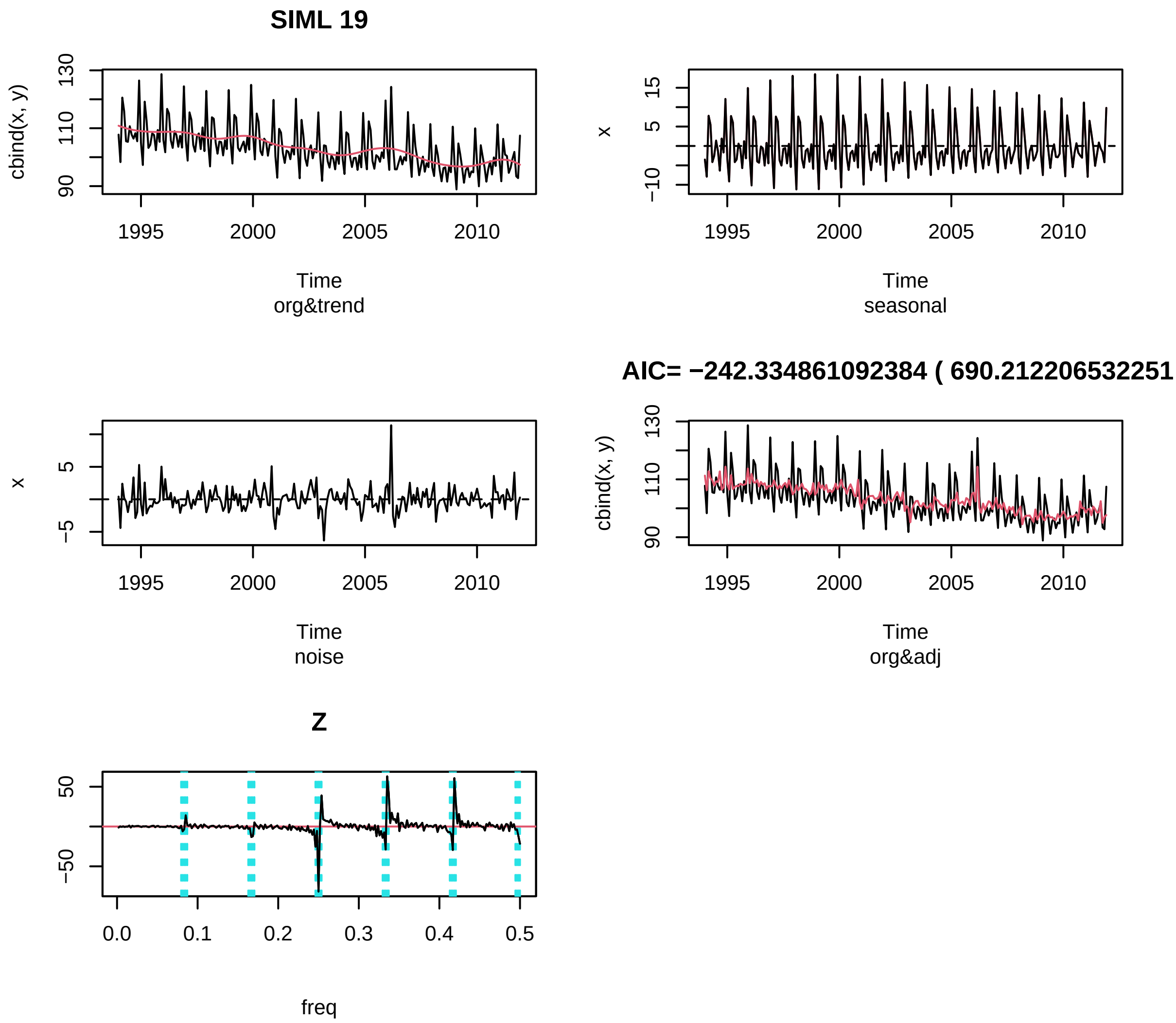

Figure 2 : S-SIML Seasonal Adjustment (Monthly Family Consumption)

problem, the proposed method is statistically simple and practically useful for seasonal adjustments in Japan and other countries.

We hope that this memorandum serves as a starting point for further developments in both the statistical issue of seasonality in economic time series and the practical applications of seasonal adjustment methods.

References

[1] Kunitomo, N. (2024), "Issues with Missing Values in Non-stationary Seasonal Economic Time Series," Unpublished Manuscript (in Japanese).

[2] Kunitomo, N. and S. Sato (2021), "A Robust-filtering Method for Noisy Non-Stationary Time Series," *Japanese Journal of Statistics and Data Science* (JJSD, Springer), 4, 373-410, Open Access.

[3] Kunitomo, N. and S. Sato (2024), *The SIML Filtering Method for Noisy Non-stationary Economic Time Series*, Springer, forthcoming.

[4] Kunitomo,N., T. Sakurai, and S. Sato (2022), "State Estimation in Economic Time Series and Macro Indicators," *Bulletin of Statistical Research*, 79, 1-20, Statistics Bureau, (in Japanese, corrected October 18, 2024): https://www.stat.go.jp/training/2kenkyu/2-2-old.html

[5] Sato, Seisho (2023), "Practical Seasonal Adjustment Using x12SIML," Web Manuscript.
(Japanese version):
https://github.com/sato-labo/x12siml/x12simldoc92.pdf
(English version):
https://github.com/sato-labo/x12siml/x12simldoc92(kuni2023-2-2).pdf

R-Data in Examples

(i) The data *shouhi* is the queterly final consumption series from Economic and Social Research Institute, Cabinet Office, JAPAN.

(ii) The data *KAKEI* is the monthly family expenditure from Statistics Bureau of JAPAN.