(In)Stability of Reg-ARIMA Models for Seasonal Adjustment

Keywords: *Seasonal Adjustment, Reg-ARIMA Models, Outliers, Leap Year Effect.*

1 Introduction

Regression models with ARIMA errors (Reg-ARIMA) are nowadays commonly used in seasonal adjustment to remove the main deterministic effects (outliers, ruptures, calendar effects) from the raw data before decomposing the corrected series into trend-cycle, seasonality and irregular.
The main seasonal adjustment programs (X-13Arima-Seats[1], Tramo-Seats[2], JDemetra+[3]) implement these models in an automatic and very user-friendly way. This facility hides in fact a real complexity and in certain cases a lack of robustness which can escape the user.
In the presentation, we draw attention to the real difficulties of implementing these models through concrete cases: the estimation of a leap year effect, the estimation of breaks and even the estimation of an ARIMA model.

This abstract focus on the case of the leap year effect. Similar results are obtainer in the case of outlier estimation.

2 Methods

2.1 Reg-ARIMA Models and Automatic Model Identification

X-13Arima-Seats and Tramo-Seats both implement a 2-step procedure to seasonally adjust a time series. In a first time, the series is corrected from deterministic effects like outliers and calendar effects using a regression model with ARIMA errors (Reg-ARIMA model) presented in Equation1. In a second time, the corrected ("linearized") series is decomposed into several components, namely the trend-cycle, the seasonality and the irregular using linear filters.

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

where $y_t'\beta$ is the regressor modeling the deterministic components of the series (constant term, outliers, calendar effects); x_t, the model residual, follows an ARIMA $(p,d,q)(P,D,Q)_s$; $\phi(B)$ and $\Theta(B)$ are polynomials in the lag operator B ($B^j z_t = z_{t-j}$); and a_t is a white noise.
A very efficient Automatique Model Identification (AMI) algorithm is implemented to automatically detect and estimate the various components of the model; see [3] for a detailed description.

[1] See [1] or [5]
[2] See [2] and [6]
[3] See [4]

2.2 Estimation of the Leap-Year effect.

The Leap-Year (LY) effect is the effect linked to the extra day periodically added to February in the calendar. This extra day is likely to have an impact on production or turnover series. The regressor associated to this LY effect is defined by:

$$LY_t = \begin{cases} 0,75 & \text{if } t \text{ is a month of February in a leap year} \\ -0,25 & \text{if } t \text{ is a month of February in a non leap year} \\ 0 & \text{otherwise} \end{cases}$$

The stability of the LY estimation is studied on a large bunch of time series using the very simple following protocol:

1) We use the monthly turnover and industrial production indexes for the members of the European Community at digits 2, 3 and 4 of the NACE revision 2 classification. This leads to a total of 2198 series.

2) In a first step, the decomposition model, the outliers, the trading-day regressors and the ARIMA model are determined on the complete time series. We use this specification in the simulations.

3) In a second step, the Reg-ARIMA is re-estimated on the 48 first observations of the series, to be sure to have one leap-year in the time span.

4) Then the estimation process is repeted adding each time a new observation which gives new estimations of the parameters. For a 13-year series, we will then have $12 \times 13 - 48 = 108$ estimations of the LY coefficient.

These simulations allow studying the convergence process of the coefficient. In this study, we assume that the LY coefficient reached convergence when the estimations remains positive, statistically significant and non statistically different from the previous estimation.

3 Results

It has to be reminded that adding an extra day to February is more than likely supposed to increase production or turnover of $1/28.25 \approx 3.5\%$ on average.
Figure 1 summarizes the time needed for the estimation to converge and Figure 2 gives more details on the convergence value.
It appears that:

- Only 624 series out of 2198 "reached convergence" (less than 30%;
- For 75% of the series that converge, 10 years at least are needed to achieve convergence.
- In median, turnover series converge to the "expected" value when IPI series converge to a larger value.
- For about 20% of the series, the convergence value seems irrealistic and sometimes larger than 10%.

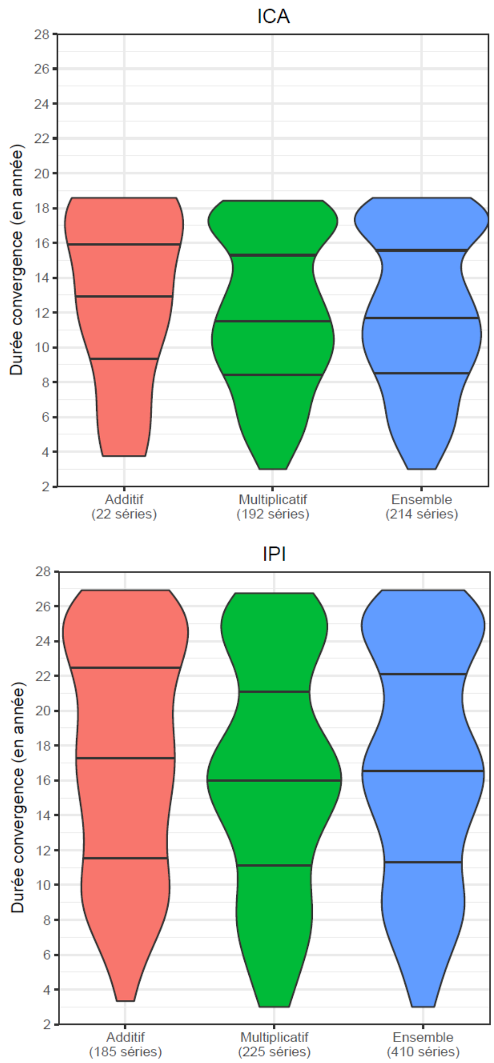

Figure 1: Convergence time (in years) of the LY coefficient (study on 624 séries).

4 Conclusions

This very simple simulation shows how instable Reg-ARIMA estimates might be. The same kind of results holds for the outliers and trading-day effects (see presentation). But these results are not really surprising. For example, to estimate a Leap Year effect, you need to observe at least 3 leap years, which means in the better case 9 years. For a “curious value”, it takes also time to see if it is an evolution of the seasonal component or an outlier which would affect the irregular.
It turns out that the specification of the Reg-ARIMA should be mainly based on expert and sectoral knowledge and is an utmost important step in seasonal adjustment. When the model is then well specified, the Reg-ARIMA methodology can be used to improve (or not) the model.
Using blindly the AMI will certainly lead sometimes in “curious”, not to say stupid, results like negative LY effect in production series.

References

[1] Findley, D.F., Monsell, B.C., Bell, W.R., Otto, M.C., Chen, B.C. (1998), New capabilities and methods of the X-12-ARIMA seasonal adjustment program, Journal of Business and Economic Statistics, 16, 127-152.

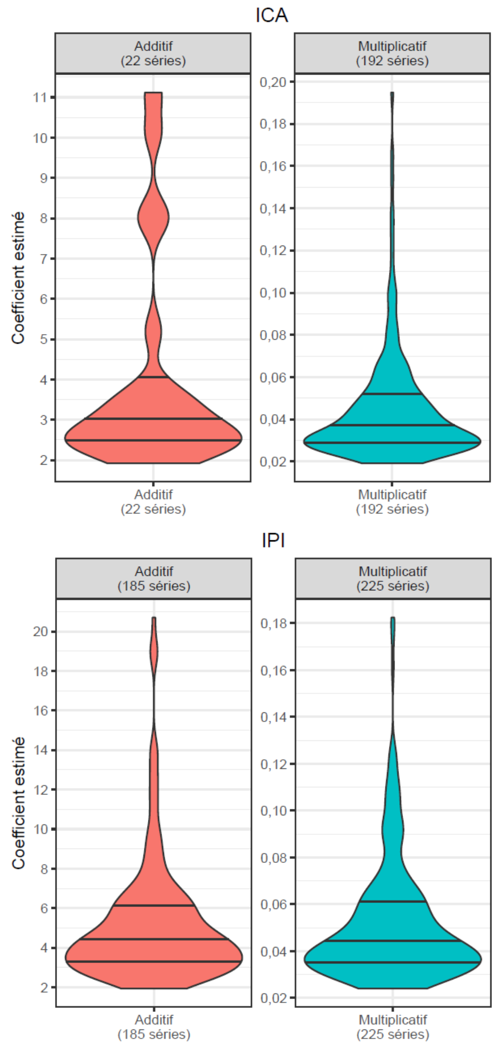

Figure 2: Final estimate of the LY effect (study on 624 séries).

[2] Gómez, V. and Maravall, A. (1996), Programs TRAMO and SEATS. Instructions for the User, (with some updates), Working Paper 9628, Servicio de Estudios, Banco de España.

[3] Gómez, V., Maravall, A. (1998), Automatic Modeling Methods For Univariate Series, Working Paper 9808, Servicio de Estudios, Banco de España.

[4] Grudkowska, S., (2017). JDemetra+ Reference Manual Version 2.2, available from URL: https://ec.europa.eu/eurostat/cros/content/software-jdemetra_en.

[5] Ladiray, D., Quenneville, B. (2001), Seasonal Adjustment with the X-11 Method, Lecture Notes in Statistics No 158, Springer, New York. Disponible en français sur https://www.census.gov/ts/papers/x11_french.pdf.

[6] Maravall, A., Caporello, G. (2004), Program TSW: Revised reference manual. Technical Report, Research Department, Bank of Spain.