Fractional and seasonal filtering

Dominique Guegan, Laurent Ferrara

Fractional and Seasonal Filtering

L. Ferrara *and D. Guégan †

Abstract

We introduce in this study a new strategy to model simultaneously persistence and seasonality inside economic data using different stochastic filters based on the Gegenbauer modelling. The limits and advantages of these filters are discussed in order to improve the adjustment of economic series, particularly when specific trend is observed. The series of new cars registrations in the Euro-zone is modelled using the previous filters.

Keywords: Persistence - Seasonality - Fractional filter - Euro-zone new car registrations.

JEL classification: C22, C52

*Banque de France, DGEI-DAMEP and CES Université Paris 1 Panthéon-Sorbonne, e-mail : ferrara@ces.ens-cachan.fr

†PSE, MSE - CES, Université Paris 1 Panthéon-Sorbonne, 106 boulevard de l'hopital, 75013 Paris, e-mail : dominique.guegan@univ-paris1.fr

1 Introduction

Many economic time series display seasonal fluctuations inherent to the economic activity. Therefore, models allowing to describe the seasonal components of the data are essential to accurately analyse and forecast business quantities. During a long time, in economic time series, the seasonal and cyclical movements have been described extending short memory processes, including seasonal components in the models, like with the classical SARIMA (Seasonal Autoregressive Integrated Moving Average) model for instance, see Box and Jenkins (1976). Those models have the specificity to exhibit peaks at seasonal frequencies in the spectral density. Now, several kinds of processes have been proposed in the literature to take simultaneously long range dependence and seasonality (see for instance Arteche and Robinson, 2000). They similarly produce peaks in the spectral density at frequencies non-necessarily equal to the seasonal ones.

In this note, we describe two alternative filters able to take into account long memory as well as seasonality. We discuss their limits and advantages for economic applications and we provide an application to new cars registrations in the Euro-zone. This approach only deals with a comparison of filters taking trend and seasonality into account in a stochastic way. In this study, we do not consider models with deterministic trend and/or seasonality, which refer to another classes of models, and we do not discuss tests and parameter estimation issues, which are other topics of research.

2 Methodology

To take into account both seasonality and long memory, various strategies can be carried out. Here, we retain two possible approaches: a rigid fractional filter and a flexible seasonal-cyclical long memory filter. The former is a natural extension of the classical SARIMA model. However, the rigidity of this model does not permit to detect several persistences characterizing the cycles. The latter one permits to bypass this issue.

2.1 Rigid fractional filters

The classical SARIMA filter, defined by $F(B) = (I - B)(I - B^s)$ has been generally used in applied economic papers dealing with seasonal time series, with $s = 12$ for monthly data or $s = 4$ for quarterly data, B being the backshift operator. This filter has been proven to be useful to make vanish trend and seasonality. However, there always exists a risk of over-differenciation implying thus a loss of information. In this respect, fractional filters avoid over-differenciation by allowing to the differenciation degree to belong to the interval $[0, 1]$.

By extending the classical SARIMA filter, Porter-Hudak (1990) has introduced the following ARFISMA (Autoregressive Fractionally Integrated Seasonally Moving Average) filter for time series with a given seasonality s:

$$F(B) = (I - B)^d (I - B^s)^D, \tag{1}$$

with $d \in R$ being the degree of integration of the long-term cycle and $D \in R$ being the degree of integration of the seasonal part. We note that for $D = d = 1$, we get the SARIMA filter. The filter (1) associates to each seasonal frequency, $\lambda_i = \frac{2\pi i}{s}$, $i = 1, \cdots, [s/2]$, the same long memory parameter D. This means that all the cycles inherent to the seasonality are characterized by the same degree of persistence. Several economic time series can be adequately described by this filter, see for example Caporale and Gil-Alana (2006) for an application on US money stock. The ARFISMA filter (1) permits also to describe the existence of an infinite cycle, for very low frequencies tending to zero, characterized by the degree of persistence $d + D$. Then, this model appears as an interesting alternative to the SARIMA approach insofar as it allows an intermediate solution to the classical choice of $d = 0$ *versus* $d = 1$, or $D = 0$ *versus* $D = 1$, encountered in unit root, or seasonal unit root, tests.

Now, it could be judicious to allow those seasonal frequencies to be characterized by different kinds of persistence. Moreover, we often get in practice estimated values for d and D describing a non-stationary behavior. It may be that only few seasonal frequencies are affected by non-stationarity, while others are not. It appears interesting to find a way to take this feature into account. We present here a filter modelling those characteristics.

2.2 Seasonal-Cyclical Long Memory (SCLM) filter

In order to allow for different persistence parameters across different frequencies, we can consider the following general representation for a seasonal long memory process. The Seasonal-Cyclical Long Memory (SCLM) filter is defined as follows:

$$F(B) = (I - B)^{d_0} \prod_{i=1}^{k-1} (I - 2B\cos\lambda_i + B^2)^{d_i} (I + B)^{d_k}, \tag{2}$$

for $i = 1, \cdots, k-1$, λ_i can be any frequency lying on the interval $]0, \pi[$ and $d_i \in R$. This filter permits to model at the same time an infinite cycle with persistence d_0 and several cycles associated to the frequencies λ_i characterized by different persistences d_i. Thus, it is possible to distinguish cycles with a stationary behavior from cycles with a non-stationary behavior, looking at the values obtained for the estimated d_i and comparing them to the value 1/2. This feature is interesting to better understand the elements characterizing the seasonal behavior of an economic series and in order to improve the forecasts of such a series, knowing that stationary elements present stronger predictive ability, Ferrara and Guégan (2001).

The filter (2) has been introduced in details by Robinson (1994) and studied by Arteche and Robinson (2000). The general representation (2) nests a lot of competitive seasonal or cyclical fractional models introduced in the literature. It includes as particular case, the k-factor Gegenbauer filter discussed by Gray, Zhang and Woodward (1989) (for $k = 1$) and then extended by Woodward *et al.* (1998) and Giraitis and Leipus (1995) whose representation is:

$$F(B) = \prod_{i=1}^{k} (I - 2\nu_i B + B^2)^{d_i}. \tag{3}$$

For $i = 1, \ldots, k$, the frequencies $\lambda_{G_i} = \cos^{-1}(\nu_i)$ being the Gegenbauer frequencies or the *G-frequencies*. Another particular case is the flexible filter introduced by Hassler (1994) for quarterly time series, factorizing the polynomial $(I - B^4)$ according to its unit roots, i.e.:

$$F(B) = (I - B)^{d_0} (I + B^2)^{d_1} (I + B)^{d_2}. \tag{4}$$

This filter allows to model stationary fractional seasonalities and here we have $k = 2$ and $\lambda_0 = 0$, $\lambda_1 = \pi/2$ and $\lambda_2 = \pi$. Now, when two fixed seasonalities exist (for example, with monthly data, $s_1 = 3$ and $s_2 = 12$) with *a priori* no infinite cycle, we can use the SFARMA (Seasonal fractionally Autoregressive Moving average) model introduced by Ray (1993). Obviously, the filter (2) contains the fractional FARMA (Fractional Autoregressive Moving Average) filter introduced by Granger and Joyeux (1980) and Hosking (1981).

2.3 Empirical remarks

From a practical point of view, in presence of a great numbers of explosions in the periodogram, even if the filters (2) and (3) seem more adequate, it can be preferable to use the filter (1) for parcimony considerations. Such an example is provided in Ferrara and Guégan (2006).

When a series presents a trend, as well as an infinite cycle and a periodic cycle of period $2\pi/\lambda$, a natural modelling would be:

$$(I - B)(I - B)^{d_0}(I - 2\cos(\lambda)B + B^2)^{d_1}X_t = \varepsilon_t. \tag{5}$$

This latter models matches the classical decomposition of macroeconomic variables. But, the long term forecasts obtained from this representation show that it is difficult to anticipate the turning points. Indeed, the influence of the linear trend remains and a long-term change in the behavior of the series is difficult to obtain.

Now a competitive approach could be the use of the following representation:

$$(I - B)^2(I - 2\cos(\lambda)B + B^2)^{d_1}X_t = \varepsilon_t. \tag{6}$$

In that latter case, the filter $(I - B)^2$ would permit, in presence of turning points, to rapidly detect them. The curve implied by this filter appears more adapted and more flexible than the use of the filter $(I - B)(I - B)^{d_0}$.

There is no statistical test able to distinguish between the two representations (5) and (6). A deep analysis of the studied series based on their main characteristics like the presence of turning points, the macroeconomic features, etc, will be useful in order to choose between these two representations, when some kind of economic trend is present. Finally, a statistical analysis of the forecasts could be the last step to retain the best model. In the next section, we illustrate some of these points.

3 Application to new passenger car registrations in the Euro-zone

Here, we consider the monthly series of new passenger car registrations in the Euro-zone released each middle of month by the Association of European Automobile Manufacturers (ACEA, see the web site www.acea.be, for further details). The raw series will be analysed from January 1960 to December 2005 (see figure 1). This series is of great interest for short-term economic analysis because it reflects, on a monthly basis, information on manufacturing goods consumption in the Euro-zone, only available on a quarterly frequency through the official quarterly accounts. Therefore, economists and market analysts follow carefully the evolution of this series, as well as the retail sales series, to have a monthly opinion of households consumption in the Euro-zone. This series is also integrated in large macroeconomic models in order to predict the Euro-zone growth (see, for example, the European Commission DG-EcFin model developed by Grassman and Keereman, 2001). However, it is well known among practitioners that, due to its high volatility, the extraction of a clear economic signal from this series is not an obvious task.

Two main stylized facts emerge strongly from this raw series, denoted $(X_t)_t$ for $t = 1, \ldots, T = 552$, namely trend and seasonality. If we look at the spectral density of the series estimated by the raw periodogram (see figure 2), we identify 7 peaks corresponding to the very low frequencies and to the seasonal frequencies $\lambda_i = 2\pi i/12$, for $i = 1, \ldots, 6$.

We compare the classical SARIMA filter and the ARFISMA fractional seasonal filter (1) with a generalized long memory approach represented by SCLM filters with $k = 6$ factors. We consider three types of parametrisation of the filter:

$$F_1(B) = (I - B)^{d_0} \Pi_{i=1}^5 (I - 2\nu_i B + B^2)^{d_i} (I + B)^{d_6}, \tag{7}$$

$$F_2(B) = (I - B)(I - B)^{d_0} \Pi_{i=1}^5 (I - 2\nu_i B + B^2)^{d_i} (I + B)^{d_6}, \tag{8}$$

$$F_3(B) = (I - B)^2 \Pi_{i=1}^5 (I - 2\nu_i B + B^2)^{d_i} (I + B)^{d_6}, \tag{9}$$

for which the frequencies are $\lambda_i = 2\pi i/12$, for $i = 1, \ldots, 5$. That is, using the filter (7) we assume that d_0 is completely unknown. Then we assume that, because of the trend, the degree of integration is of the form $1 + d_0$ and we use the filter (8). Finally, we assume the existence of an integrated process of order 2, that is $d_0 = 2$ and we use the filter (9).

We compare the accuracy of the models through their goodness of fit by using the Ljung-Box $(Q(K))$ and Jarque-Bera (JB) statistics on the filtered series. These statistics are presented in table 1. The estimation step for models is reduced to the estimation of the memory parameters. In this approach, we do not consider the short memory behaviour of the series, we are rather interested in modelling the medium or long-term macroeconomic component of the series. In the following, the noise $(\varepsilon_t)_t$ is assumed to be an ergodic noise (for a prediction approach, it can be a strong white noise, an ARMA process or a kind of GARCH process).

Here, the Whittle Pseudo-Maximum-Likelihood method is carried out to estimate those parameters. We do not use the semi-parametric method developed by Robinson (1994), nor the local discrete version of the Whittle quasi-likelihood function (Arteche, 2003). The simplex algorithm is used first to determine more precisely the initial values, then the BHHH algorithm is employed. Parameter estimates of the models are presented in table 2. They are all significatively different from zero, but they imply a non-stationarity in almost all the frequencies. When comparing the ARFISMA filter with SCLM filters, we observe that the sum of $\hat{d}$ and $\hat{D}$ is close to $\hat{d}_0$ in F_1 and F_2, illustrating that the long-term persistence in ARFISMA is characterized by both parameters, although it is not possible to distinguish their influence for very low frequencies. This fact also implies that $\hat{D}$ is lower than $\hat{d}_i$ for all seasonal frequencies underlying the rigidity of the ARFISMA model. When comparing now the parameters of both filters F_1 and F_2, we note that the degree of persistence for very low frequencies is similar (1.2072 for F_1 and 1.1284 for F_2), but the seasonal memory parameters for F_2 are all greater. This phenomena points out the strong influence of low frequencies persistence in the data, that dominates the seasonal persistence. Comparing all the filters through the results given in table 1, we observe that the filter F_2 provides the smallest residual variance implying thus a better fit to the data. Now, the Jarque-Bera statistics is also the smallest for this filter, albeit significantly different from zero at the usual type I risks, implying thus non-Gaussianity. The reduction of the JB statistics is equally due to a reduction of both skewness and kurtosis for this last model. Lastly, the Llung-Box statistics, $Q(K)$, is also strongly reduced for each K using the filter F_2. Consequently we conclude that this latter filter improves all the criteria and removes correctly both the seasonality and the persistence observed in the data. This little empirical experience shows that it is more convenient to first differentiate data with a strongly marked trend and then to estimate the memory parameters. Moreover, we have again the confirmation that a degree of integration equal to 2 (filter F_3) is too strong to improve the goodness of fit of macroeconomic data, although the forecasting results may be improved in terms of mean-squared-error. Indeed, filter F_3 over-differentiates the data, implying a strong increase of all the seasonal memory parameters, one of them being greater than one.

With this example we show the interest to use competitively those different filters in the simultaneous presence of persistence and long memory inside economic data. It is noteworthy that the use of a seasonal cyclical long memory filters does not necessitate a great amount of data (around 10 years of data are enough), we refer for instance to studies using Monte Carlo simulations, see Ferrara (2000) and Ferrara and Guégan (2001).

4 Conclusion

Many economic time series display persistent seasonal fluctuations, but it does not exists a single way to model the seasonal components and the long-term cycle. In this paper we recall most of the seasonal fractionally integrated processes which model trend and seasonality. We show on an application to economic activity in the Euro-zone that seasonal-cyclical long memory filters offer very competitive alternatives to the classical linear SARIMA filter. A forecast study on various macroeconomic series, allowing to compare the prediction ability of the various non-linear cyclical long memory models, is the subject of a companion paper. As a further extension of this work, it would be interesting to extend this approach to the multivariate framework in order to take into account the common evolution of long cycles and seasonality, present in many economic series.

References

[1] Arteche, J., 2003, Semiparametric robust tests on seasonal or cyclical long memory time series, Journal of Time Series Analysis 23, 251-285.

[2] Arteche, J. and P.M. Robinson, 2000, Semiparametric inference in seasonal and cyclical long memory processes, Journal of Time Series Analysis 21, 1-25.

[3] Box, G.E.P. and G.M. Jenkins, 1976, Time series analysis: Forecasting and control, Holden Day, San Francisco.

[4] Caporale, G.M., and L.A. Gil-Alana, 2006, Long memory at the long-run and the seasonal monthly frequencies in the US money stock, Applied Economics Letters 13, 965-968.

[5] Ferrara, L., 2000, Processus Longue Mémoire Généralisés: Estimation, Prévisions et Applications, PhD Thesis, University Paris XIII. France.

[6] Ferrara, L. and D. Guégan, 2001, Forecasting with k-factor Gegenbauer processes: Theory and applications, Journal of Forecasting 20, 581-601.

[7] Ferrara, L. and D. Guégan, 2006, Fractional seasonality: Models and applications to economic activity in the euro area, in EuroIndicators Papers Collection, Eurostat Publications, Luxembourg.

[8] Giraitis, L. and R. Leipus, R., 1995, A generalized fractionally differencing approach in long memory modeling, Lithuanian Mathematical Journal 35, 65-81.

[9] Granger C.W.J. and R. Joyeux, 1980, An introduction to long memory time series models and fractional differencing, Journal of Time Series Analysis 1, 15-29.

[10] Grassman, P. and F. Keereman, 2001, An indicator-based short-term forecast for quarterly GDP in the euro area, Economic Paper 154, European Commission.

[11] Gray, H.L., N.-F. Zhang and W.A Woodward, 1989, On generalized fractional processes, Journal of Time Series Analysis 10, 233-257.

[12] Hassler, U., 1994, Misspecification of long memory seasonal time series, Journal of Time Series Analysis 15, 19-30.

[13] Hosking, J.R.M., 1981, Fractional differencing, Biometrika 68, 1, 165-176.

[14] Porter-Hudak, S., 1990, An application to the seasonally fractionally differenced model to the monetary aggregates, Journal of the American Statistical Association 85, 338-344.

[15] Ray, B.K., 1993, Long-range forecasting of IBM product revenues using a seasonal fractionally differenced ARMA model, International Journal of Forecasting 9, 255-269.

[16] Robinson P.M. (1994) Efficient tests of non-stationary hypotheses, Journal of the American Statistical Association, 89, 1420 -1457.

[17] Woodward, W.A., Q.C. Cheng and H.L. Gray, 1998, A k-factor GARMA long-memory model, Journal of Time Series Analysis 19, 5, 485-504.

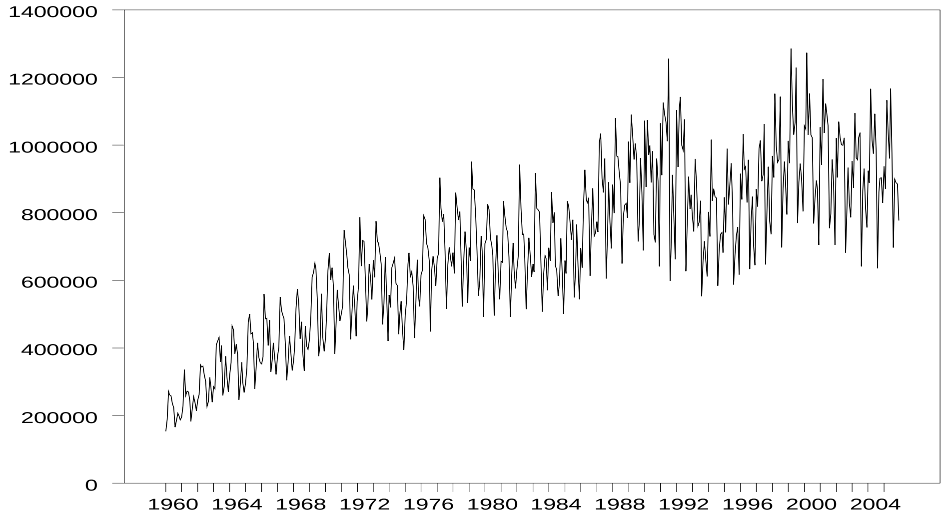

Figure 1: New car registrations in the Euro-zone from Jan. 1960 to Dec. 2005 (raw series, source ACEA.)

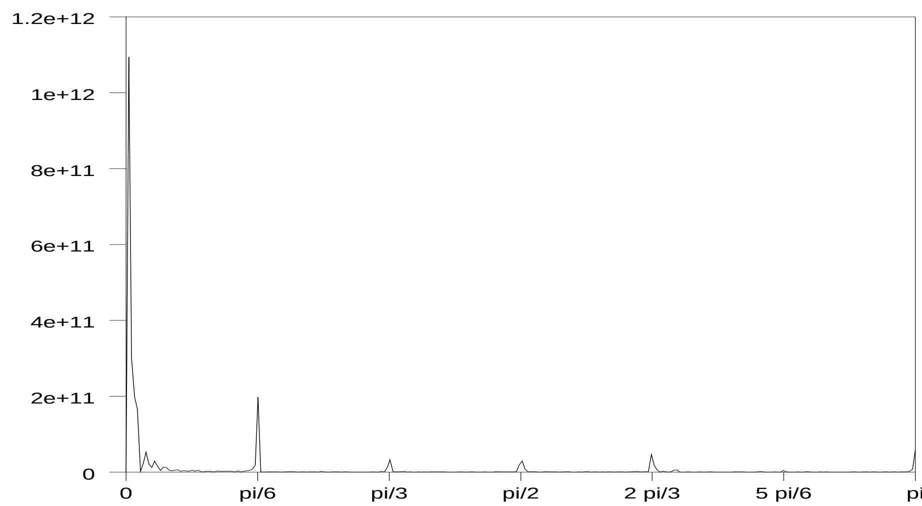

Figure 2: Raw periodogram of the new car registrations series in the Euro-zone from Jan. 1960 to Dec. 2005.

Filter F(B)	$\hat{d}$	$\hat{D}$	$\hat{\sigma}_\varepsilon$	JB stat	Q(12)	Q(24)	Q(60)	Q(120)
$(I-B)(I-B^s)$			78689	522	185	340	696	1362
$(I-B)^d(I-B^s)^D$	0.6892	0.5894	62579	474	139	284	635	1332
$F_1(B)$			58713	396	30	97	278	632
$F_2(B))$			57044	306	14	69	212	481
$F_3(B)$			63822	433	44	103	263	533

Table 1: Analysis of SARIMA, ARFISMA and SCLM filters applied to the new car registrations series in the Euro-zone from Jan. 1960 to Dec. 2005, using empirical variances, JB and Portmanteau statistics. Estimation of the parameters d, D, for the ARFISMA filter.

	Filters	$F_1(B)$	$F_2(B)$	$(F_3(B)$
i	$\hat{\lambda}_i$	$\hat{d}_i$	$\hat{d}_i$	$\hat{d}_i$
0	0.0114	1.2072	0.1284	
1	0.5255	0.5725	0.6481	0.5851
2	1.0510	0.4730	0.5381	0.6637
3	1.5765	0.5451	0.6088	0.7438
4	2.0906	0.7474	0.8336	1.0190
5	2.6161	0.4465	0.5193	0.6998
6	3.1416	0.6608	0.7208	0.8738

Table 2: Parameter estimation of parameters of the three SCLM filters introduced respectively in (7), (8) and (9), for the new car registrations series in the Euro-zone from Jan. 1960 to Dec. 2005.