Seasonalities and cycles in time series: A fresh look with computer experiments

Michel Fliess, Cédric Join

To cite this version:


HAL Id: hal-01208171

https://hal-polytechnique.archives-ouvertes.fr/hal-01208171

Submitted on 2 Oct 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Seasonalities and cycles in time series: A fresh look with computer experiments

Michel Fliess*∗∗∗∗∗∗, Cédric Join *∗∗∗∗∗∗∗∗∗∗∗∗

LIX (CNRS, UMR 7161), École polytechnique, 91128 Palaiseau, France (e-mail: Michel.Fliess@polytechnique.edu)

CRAN (CNRS, UMR 7039), Université de Lorraine, BP 239, 54506 Vandœuvre-lès-Nancy, France (e-mail: cedric.join@univ-lorraine.fr)

AL.I.E.N. (ALgèbre pour Identification & Estimation Numériques), 24-30 rue Lionnois, BP 60120, 54003 Nancy, France (e-mail: {michel.fliess, cedric.join}@alien-sas.com)

Projet Non-A, INRIA Lille – Nord-Europe, France

Abstract: Recent advances in the understanding of time series permit to clarify seasonalities and cycles, which might be rather obscure in today’s literature. A theorem due to P. Cartier and Y. Perrin, which was published only recently, in 1995, and several time scales yield, perhaps for the first time, a clear-cut definition of seasonalities and cycles. Their detection and their extraction, moreover, become easy to implement. Several computer experiments with concrete data from various fields are presented and discussed. The conclusion suggests the application of this approach to the debatable Kondratiev waves.

Keywords: Time series, seasonalities, cycles, decomposition, detection, extraction, trend, deseasonalization, Kondratiev waves, nonstandard analysis, time scales.

1. INTRODUCTION

Everyone who is interested in econometrics, finance, sales, various questions related to management, and in many other topics, like, for instance, the study of health and climate data (see, e.g., Barnett & Dobson (2010); Mudelsee (2010)), has heard the words seasonalities and cycles, as well as the word deseasonalization. Although seasonalities, cycles, their detection and their extraction may be found since ever in many publications (see, e.g., Baum & Lundtorp (2001); Bourbonnais & Terraza (2010); Geman (2015); Ghysels & Osborn (2001); Holtfort (2009); Levin & Rubin (1996); Mélard (2008); Tsay (2010), and the references therein), this subject, which plays such an fundamental rôle in not only concrete but also abstract questioning, seems to be far from being fully understood (see, e.g., Bell, Holan & McElroy (2012)). A lack of any clear-cut definition might be the explanation. Seasonalities and cycles are seldom strictly periodic phenomena (see, e.g., Bourbonnais & Terraza (2010); Franses & Paap (2004)), which could be then detected and extracted thanks to classic means like Fourier techniques stemming from signal processing. The actual situation which might be summarized from today’s literature is rather obscure:

• Seasonalities are often, but not always, related to time horizons which are shorter than one year. Time horizons for cycles are often, but not always, longer.
• Seasonalities are often, but not always, assumed to exhibit a “more” periodic behavior than cycles.

Tedious ad hoc approaches are therefore developed for concrete case-studies (see, e.g., Borovkova & Geman (2006), Hamilton (1989), Barnett & Dobson (2010), Jönsson & Eklundh (2002), Koopman & Lee (2009), Zhang & Qi (2005), several contributions in Bell, Holan & McElroy (2012), and the references therein).

The following additive decomposition is often encountered (see, e.g., Bourbonnais & Terraza (2010); Brockwell & Davis (1991); Ghysels & Osborn (2001); Mélard (2008)): $µ_t + S_t + C_t + R_t$ (1)

where

(1) $µ_t$ is the trend,
(2) $S_t$ is the seasonal component,
(3) $C_t$ is the cyclic component,
(4) $R_t$ is the residual component.

Cycles are sometimes ignored. It yields to $µ_t + S_t + R_t$ (2)

From a mathematical standpoint and to the best of our knowledge, the decompositions (1) and (2) have never been rigorously proved.

Remark 1. This is confirmed by the fact that

• mathematically oriented treatises on time series, like for instance those by Hamilton (1994) and by Gourieroux & Monfort (1995),
• the historical and epistemological study due to Meurant (2012)

do not mention seasonalities and cycles in spite of their indubitable practical usefulness.
Our aim in this communication is to suggest another route and to illustrate it via some computer experiments. Two main ingredients are employed:

(1) Tools stemming from Robinson’s nonstandard analysis (see, e.g., Diener & Reeb (1989); Nelson (1987); Robinson (1996)) are used for a new theory of time series (see Fließ & Join (2009, 2015); Fließ, Join & Hatt (2011), and the references therein). A rather recent theorem due to Cartier & Perrin (1995) yields a new understanding of their structure.

(2) Several time scales.

Detection and extraction of seasonalities and cycles become much simpler. The quite involved techniques, like nonstationary time series modeling, which are often employed in the literature, become pointless in our model-free setting (see also Fließ & Join (2013) for further comments on this viewpoint).

Our communication is organized as follows. Section 2 is sketching our approach to time series. Seasonalities, cycles and deseasonalization are defined in Section 3. Section 4 shows and discusses computer experiments. Let us emphasize the last example, in Section 4.4, where the introduction of several time scales, which was suggested in Section 3.1, turns out to be most enlightening. Some concluding comments, with a hint on the celebrated but controversial Kondratiev waves, may be found in Section 5.

2. REVISITING TIME SERIES

2.1 Time series and nonstandard analysis

Take a time interval $[0,1]$. Introduce as often in nonstandard analysis the infinitesimal sampling

$$T = \{0 = t_0 < t_1 < \cdots < t_\nu = 1\}$$

where $t_{i+1} - t_i$, $0 \leq i < \nu$, is infinitesimal, i.e., “very small.” A time series $X$ is a function $T \to \mathbb{R}$.

Remark 2. Infinitely small or large numbers should be understood as mathematical idealizations. In practice a time lapse of 1 second (resp. hour) should be viewed as quite small when compared to 1 hour (resp. month).

Nonstandard analysis may therefore be applied in concrete setting (see also Fliess & Join (2013) for further comments (1)

Our communication is organized as follows. Section 2 is sketching our approach to time series. Seasonalities, cycles and deseasonalization are defined in Section 3. Section 4 shows and discusses computer experiments. Let us emphasize the last example, in Section 4.4, where the introduction of several time scales, which was suggested in Section 3.1, turns out to be most enlightening. Some concluding comments, with a hint on the celebrated but controversial Kondratiev waves, may be found in Section 5.

A time series $X : T \to \mathbb{R}$ is said to be quickly fluctuating, or oscillating around 0, if, and only if, its Lebesgue-integral

$$\int_A X \, dm$$

(see Cartier & Perrin (1995)) is infinitesimal for any appreciable interval $A$, i.e., an interval which is neither infinitely small nor infinitely large.

According to a theorem due to Cartier & Perrin (1995) the following additive decomposition holds for any time series $X$, which satisfies a weak integrability condition,

$$X(t) = E(X)(t) + X_{\text{fluctuation}}(t)$$

where

- the mean $E(X)(t)$ is “quite smooth,”
- $X_{\text{fluctuation}}(t)$ is quickly fluctuating.

Decomposition (4), which is unique up to an additive infinitesimal number, is replacing here Equations (1) and (2).

Remark 3. The very definition of quick fluctuation shows that the mean $E(X)$ in Equation (4) may be estimated via moving average techniques, which are already popular in the time series literature.

3. SEASONALITIES AND CYCLES

3.1 Several time scales

From Remark 2, taking different time scales becomes obvious. Assume for instance that the time series $X(t)$ is defined via a time scale where $t_{i+1} - t_i = 1$ second in the sampling (3). The two time series

$$X(t) = \sum_{i=1}^{i=3600} X(t - \iota)$$

and

$$X(t) = \sum_{i=1}^{i=3600} \frac{X(t - \iota)}{3600}$$

are then defined with respect to a 1 hour sampling.

3.2 Some definitions

Seasonalities and cycles are defined with respect to a rather “coarse” sampling (3). Let us emphasize that different time scales might lead to different seasonalities and cycles (see Section 4.4).

Rewrite Equation (4)

$$X_{\text{fluctuation}}(t) = X(t) - E(X)(t)$$

Introduce a threshold $\varpi > 0$ around the mean $E(X)$. A crossing time $T$ is defined by the two following properties:

(1) $|X_{\text{fluctuation}}(T)| > \varpi$.

(2) $|X_{\text{fluctuation}}(T - h)| \leq \varpi$, where $h > 0$ is a sampling period.

Let $\{T_1, \ldots, T_N\}$ be the set of crossing times where, for instance, $X_{\text{fluctuation}}(T) > \varpi$. If $T_{i+1} - T_i$ is “approximately” periodic, then we have a strong seasonality, or a strong cycle. When there is no approximate periodicity, we have a weak seasonality, or a weak cycle.

3.3 Deseasonalization

A deseasonalization of a time series $X$ is nothing else than replacing it by its mean $E(X)$ according to Equation (4). In other words fluctuations are removed.

Also called average or trend. Let us emphasize that the notion of trend in the classic literature on time series is very different (see, e.g., Tsay (2010)).

Introducing a threshold is an obvious necessity for defining this notion of approximate periodicity.

We speak of seasonality if $T_{i+1} - T_i$ is less than 1 year, of cycle if not.
4. COMPUTER ILLUSTRATIONS

Data for the first three examples are provided by the French Insee (Institut national de la statistique et des études économiques). Consider three types of monthly data:

- consumer prices of several foodstuffs in France,
- job seekers registered at the French employment agency (Pôle emploi),
- international passengers in Paris airports.

The internet traffic data of the last example, in Section 4.4, are borrowed from Cortez, Rio, Rocha & Sousa (2012).

4.1 Foodstuffs prices

Select the following monthly data:

- Figure 1: apples from January 1998 to April 2015,
- Figure 2: tomatoes from January 1998 to April 2015,
- Figure 3: lambs from January 1992 to April 2015.

Equation (4) is illustrated by the Figures 1-(a), 2-(a) and 3-(a). The means are computed via a centered moving average of length 20. Figures 1-(b), 2-(b) and 3-(b) display the fluctuations. Strong seasonalities appear after some time in Figure 2-(c). There are only weak seasonalities in Figure 3-(c).

4.2 Unemployment

Figure 4-(a) displays the number of job-seekers from December 1995 to June 2015. The mean, which is estimated via a sliding window of length 20 months, and the threshold are shown in Figure 4-(b). A strong seasonality are obvious in Figure 4-(c).

4.3 Air passengers

Monthly data from January 1982 to April 2015 are presented in Figure 5-(a). The mean is obtained via a centered moving average of length 10. Figure 5-(b) shows that the magnitude of fluctuations is related to the number of passengers. \( \frac{\Delta X(t)}{\Delta X(0)} = \frac{X(t) - X(0)}{X(0)} \) is therefore presented in Figure 5-(c) as well as the threshold. Figure 5-(d) display a clear-cut strong seasonality.

4.4 Internet traffic

Figure 6 shows internet traffic data, in bits, which were collected every 5 minutes from 06:57 a.m., 7 June 2005, to 11:17 a.m., 31 July 2005. Several time scales are used via Equation (5). Figures 7-(b) and 7-(c) display respectively the fluctuations and the seasonalities with respect to a time scale of 1 hour. The trend in Figure 7-(a) is computed via a moving average of 30 points. Figure 8 presents the same type of results with a time scale of 1 day.

It should be clear that the above seasonalities have been detected and extracted only thanks to several time scales.

5. CONCLUSION

The above encouraging results need of course to

- be confirmed by several other concrete case-studies,
- yield forecasting improvements.

It would be also most rewarding to see if our viewpoint may be helpful for investigating some complex questions like, for instance, the famous Kondratieff waves (Kondratjew (1926)\(^7\)), which are quite debatable (see, e.g., Bosserelle (1994)). Are the time series which are at our disposal long enough to start such a research? A recent paper by Korotayev & Tsirel (2010), where techniques stemming from spectral analysis are used, looks quite stimulating.

REFERENCES


6 http://www.insee.fr/fr/bases-de-donnees/

7 Kondratjew is a German spelling of this Russian name.
Fig. 1. Apples (1 kg)


(a) Consumer prices in France (–blue) and its mean (–red)
(b) Fluctuations (–blue) and threshold (–red)
(c) Time lapse between fluctuations

Fig. 2. Tomatoes (1 kg)

Fig. 3. Lamb - middle neck (1 kg)
Fig. 4. Unemployment
Fig. 5. Air passengers

Fig. 6. Internet traffic per 5 minutes (original series)
Fig. 7. Internet traffic with a time scale of 1 hour
Fig. 8. Internet traffic with a time scale of 1 day