New Results on Linear Filters Minimizing Phase-Shift for Seasonal Adjustment

Fabien Guggemos* Dominique Ladiray† Michel Grun-Rehomme‡

Abstract
Moving averages or linear filters are ubiquitous in seasonal adjustment and business cycle extraction methods. For example, software packages like X-12-Arima or Tramo-Seats use Henderson's and composite moving averages to estimate the main components of a time series, notably by accounting for forecasts obtained from an ARIMA modelling. When estimating the most recent points, all these methods must rely on asymmetric filters, whose main drawback is to induce phase shift effects that usually impact the real-time estimation of turning points. Here, we first propose a timeliness criterion included in a global procedure to compute moving averages minimizing the phase shift. Second, by drawing up a general unifying framework, we get a theoretical link between this way of designing moving averages, the filters based on minimized revisions purposes and a recently developed data-driven procedure called the Generalized Direct Filter Approach. Consequently, we also show that the results obtained for moving averages can be easily extended to any kind of linear filters, like the Hodrick-Prescott filter. Empirical results and comparisons on real time series data will be eventually presented.

Key Words: Time Series, Seasonal Adjustment, Linear Filters, Moving Averages, Phase-Shift.

1. Introduction

Moving averages or linear filters are ubiquitous in seasonal adjustment and business cycle extraction methods. For example, the commonly used software package X-12-Arima uses Henderson's and composite moving averages to estimate the main components of a time series. Symmetric filters are applied to the center of the series, but when it comes to the estimation of the most recent points, all these methods must rely on asymmetric filters. X-12-Arima or Tramo-Seats apply symmetric averages on forecasts obtained from an ARIMA modeling of the series. As forecasted values are linear combinations of past values, it turns out that these methods use asymmetric moving averages at the end of the series.

If these asymmetric moving averages have good properties regarding the size of future revisions induced by the smoothing process (see [2, 3, 8]), they also induce phase shifts that usually impact the real-time estimation of turning points. Figure 1 gives an illustration of this phase shift problem. The thick black line represents the "true value" of the trend-cycle of the French industrial production index estimated in June 2010 with X-12-Arima. A turning point can be observed in January or February 2001. The colored lines show the successive estimates of the trend-cycle obtained with X-12-Arima using data up to March 2001, April 2001 .... October 2001. It is only in June 2001, and in fact 4 or 5 months after the real date, that the turning point can be noticed.

In this paper, we propose a timeliness criterion which allows us to compute moving averages minimizing the phase shift. By drawing up a general unifying framework, we get a theoretical link between this way of designing moving averages, the filters based on minimized revisions purposes and a recently developed data-driven procedure called the

*INSEE, 18 boulevard Adolphe Pinard, 75014, Paris, France
†INSEE, 18 boulevard Adolphe Pinard, 75014, Paris, France
‡ENSAE, 15 avenue Pierre Larousse, 92240, Malakoff, France

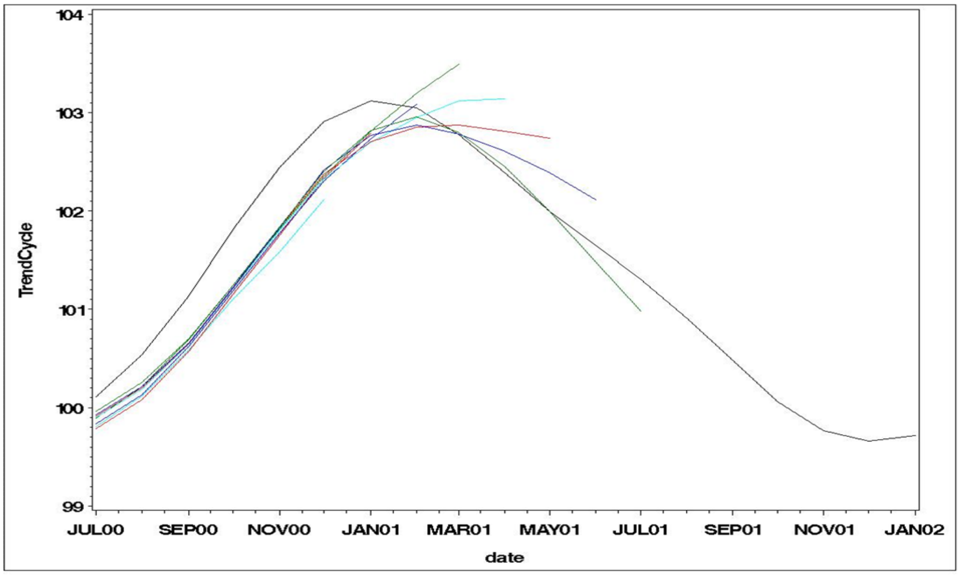

Figure 1: Real-time estimation with `X-12-Arima`.

The thick black line represents the "true value" of the trend-cycle of the French industrial production index estimated in June 2010 with X-12-Arima. The colored lines show the successive estimates of the trend-cycle obtained with X-12-Arima using data up to March 2001, April 2001 .... October 2001.

Generalized Direct Filter Approach (see [11, 12, 13]). Consequently, even if we focus here on moving averages whose coefficients do not depend on the characteristics of the series, the results can be easily extended to any kind of moving averages.

Section 2 gives some definitions, identifies mathematically both main effects (gain and phase shift) of linear filters and shows how moving averages preserving some trends or cancelling some seasonalities can be designed. Section 3 presents the revision criterion introduced by John Musgrave in 1964 and draws a parallel with the Wildi's Direct Filter Approach, the generalization of which emphasizes the necessity of minimizing the phase shift. Section 4 then shows that the construction of such linear filters can be seen as particular cases of a same very general unifying framework. By varying some modeling assumptions, it provides a large class of linear filters, data independent as well as data dependent, including not only those based on revisions, but also some well-known moving averages like Henderson's ones. The generalized Direct Filter Approach then yields an interesting point of comparison to the new Timeliness criterion introduced in Section 5, that permits to minimize the phase shift effects. Finally, in Section 6, we present the global procedure to construct asymmetric moving averages that induce a phase shift as small as possible, followed by some empirical results.

2. Moving averages: definitions and design

From now on, let $(X_t)_{t\in\mathbb{Z}}$ denote a time series that can be usually considered as the sum of three components - the trend-cycle TC_t, the seasonality S_t and the irregular component I_t - with $X_t = TC_t + S_t + I_t$. In case of a multiplicative model, the logarithmic transformation of $(X_t)_{t\in\mathbb{Z}}$ should be rather considered. Moving averages are the basic tool of the X-12-ARIMA seasonal adjustment method and are used to estimate these three main components[1].

[1] In X-12-Arima and Tramo-Seats, the other components, namely the outliers and the trading-day effects, are detected and estimated before the seasonal adjustment procedure per se (see [7] for a detailed presentation

2.1 Definitions

Let p and f be two nonnegative integers and $\boldsymbol{\theta} = (\theta_{-p}, \ldots, \theta_{+f})'$ a vector of $p+f+1$ real numbers. The moving average $L_{\boldsymbol{\theta}}$ is then defined as the linear endomorphism of the vector space of time series which associates to every series $(X_t)_{t\in\mathbb{Z}}$, also called the input signal, the series $(Y_t)_{t\in\mathbb{Z}} = (L_{\boldsymbol{\theta}} X_t)_{t\in\mathbb{Z}}$, also called the output signal, such that:

$$\forall t \in \mathbb{Z},\ Y_t = L_{\boldsymbol{\theta}} X_t = \sum_{k=-p}^{+f} \theta_k X_{t+k}.$$

The value at instant t of the unadjusted series is therefore replaced by a weighted average of p "past" values of the series, the current value, and f "future" values of the series.

- The quantity $p + f + 1$ is called the **moving average order**.
- When p is equal to f, that is, when the number of points in the past is the same as the number of points in the future, the moving average is said to be **centered**.
- If, in addition, $\theta_{-k} = \theta_k$ for any k, the moving average is said to be **symmetric**.

Generally, the observations X_t are available only for a finite set of indices $t = 1, \ldots, T$. Consequently, with a moving average of order $p + f + 1$ calculated at instant t with p points in the past and f points in the future, it will be impossible to smooth out the first p values and the last f values of the series. In the X-11 method, symmetric moving averages play an important role; but, to avoid losing information at the ends of the series, they are supplemented with ad hoc asymmetric moving averages.

If we want for example to get an estimation of the trend-cycle from the decomposition model $X_t = TC_t + S_t + I_t$ using the moving average $L_{\boldsymbol{\theta}}$, it should be desirable that $L_{\boldsymbol{\theta}}$ remove the seasonality S_t, reduce the irregular component I_t as much as possible and preserve the trend-cycle TC_t.

2.2 Preserving trends and removing seasonalities

It is possible to design a moving average, i.e. determining the set of coefficients $\{\theta_k\}$, that preserves simple trends, particularly polynomials. For instance, for any moving average to preserve constant series $X_t = a$, it is necessary that $a = L_{\boldsymbol{\theta}} X_t = \sum_k \theta_k X_{t+k} = a\cdot(\sum_k \theta_k)$, and therefore that the sum of the coefficients of the moving average, $\sum_{k=-p}^{+f} \theta_k$, be equal to 1. When looking at preservation of straight lines $X_t = at + b$, it is necessary that, for any t, $at+b = L_{\boldsymbol{\theta}} X_t = \sum_k \theta_k \left[a(t+k) + b\right] = (at+b)\cdot(\sum_k \theta_k) + a\cdot(\sum_k k\theta_k)$, which results in $\sum_{k=-p}^{+f} \theta_k = 1$ and $\sum_{k=-p}^{+f} k\theta_k = 0$. Generally, it can be shown that for a moving average to preserve a polynomial of degree d, it is necessary and sufficient that its coefficients satisfy the following conditions:

$$\sum_{k=-p}^{+f} \theta_k = 1 \ \text{ and } \ \sum_{k=-p}^{+f} k^j \theta_k = 0,\ j = 1, \ldots, d.$$

It is also possible to design a moving average that cancels a seasonality which can evolve polynomially with time. Let us note $S_t = (a_0 + a_1 t^1 + a_2 t^2 + \ldots + a_d t^d) u_t$ where u_t is a periodical function with period ℓ. And let us define n the integer satisfying $p + f + 1 = n\ell$ where $p + f + 1$ is the order of the moving average.

of the X-12-Arima algorithm).

It can be shown[2] that for a moving average to cancel a seasonality evolving like a polynomial of degree d, it is necessary and sufficient that its coefficients satisfy:

$$\begin{cases} \sum_{j=0}^{n-1} \theta_{k-j\ell} - \sum_{j=0}^{n-1} \theta_{f-j\ell} & = & 0 \\ \sum_{j=0}^{n-1} (k-j\ell)\theta_{k-j\ell} - \sum_{j=0}^{n-1} (f-j\ell)\theta_{f-j\ell} & = & 0 \\ \dots & & \\ \sum_{j=0}^{n-1} (k-j\ell)^d\theta_{k-j\ell} - \sum_{j=0}^{n-1} (f-j\ell)^d\theta_{f-j\ell} & = & 0 \end{cases}$$

Preserving trends and removing seasonalities imply therefore that the choice of coefficients is generally subject to some minimal linear constraints. More generally, the subset of admissible vectors $\boldsymbol{\theta}$ in $\mathbb{R}^{p+f+1}$ - i.e. the subset of vectors satisfying every constraint - is supposed to be **convex** and will be denoted by Θ.

2.3 Gain and phase shift effects

To identify both main effects of the moving averages, let us now consider an harmonic time series at frequency ω, $X_t(\omega) = e^{i\omega t}$. The transform of X_t by any moving average $L_{\boldsymbol{\theta}}$ will be:

$$Y_t = L_{\boldsymbol{\theta}} X_t = \sum_{k=-p}^{+f} \theta_k X_{t+k} = \sum_{k=-p}^{+f} \theta_k e^{i\omega(t+k)} = \left(\sum_{k=-p}^{+f} \theta_k e^{i\omega k} \right) \cdot X_t$$

Mathematically, the harmonic time series appear as the eigenvectors of operator $L_{\boldsymbol{\theta}}$, whose eigenvalues are respectively the values taken by the transfer function of the filter, $\widehat{L}_{\boldsymbol{\theta}}\left(e^{i\omega}\right) = \sum_{k=-p}^{+f} \theta_k e^{i\omega k}$.

If, for any frequency ω, $\rho_{\boldsymbol{\theta}}(\omega)$, $\varphi_{\boldsymbol{\theta}}(\omega)$ denote respectively the modulus and the argument in $]-\pi;\pi]$ of this transfer function, then we have $L_{\boldsymbol{\theta}} e^{i\omega t} = \rho_{\boldsymbol{\theta}}(\omega)\, e^{i[\omega t + \varphi_{\boldsymbol{\theta}}(\omega)]}$, or, to stay within real time series domain,

$$L_{\boldsymbol{\theta}} \cos(\omega t) = \rho_{\boldsymbol{\theta}}(\omega) \cos\left[\omega t + \varphi_{\boldsymbol{\theta}}(\omega)\right].$$

Applying a moving average to an harmonic time series then transform it in two different ways : on the one hand, by multiplying it by an amplitude coefficient $\rho_{\boldsymbol{\theta}}(\omega)$; on the other hand, by 'shifting' it in time by $\varphi_{\boldsymbol{\theta}}(\omega)/\omega$, which directly affects the detection of turning points. In signal theory, $\rho_{\boldsymbol{\theta}}(\omega)$ and $\varphi_{\boldsymbol{\theta}}(\omega)$ are simply the gain and the phase of the transfer function $\widehat{L}_{\boldsymbol{\theta}}\left(e^{i\omega}\right)$. For any stationary time series $(X_t)_{t\in\mathbb{Z}}$, its component at frequency ω is then affected by these two distinct effects.

3. Optimal moving averages based on revision criteria

3.1 Musgrave's approach

The filters commonly used for seasonal adjustment are derived by tuning their characteristics to application purposes, like minimization of revisions. In particular, Doherty ([2]) reviewed the method of Musgrave ([8]) for calculating the X-12-Arima trend-cycle asymmetric averages. A summary of this method is given hereinafter.

Let $\{w_1, \dots, w_N\}$ with $N = 2p+1$ be the weights of the Henderson's moving average, the definition of which is recalled in section 4.2. Let also $\{\theta_1, \dots, \theta_M\}$ $-$ with $p+1 \leq M \leq 2p$ and $\sum_{i=1}^{M} \theta_i = 1$ $-$ be one of the p vectors of Musgrave's asymmetric averages.

[2] see [4] for a complete proof

There are p such averages, one for each of the last p points. Assume that M is fixed and that the last data points of the time series follow a simple linear trend of the following form,

$$X_t = a + bt + \varepsilon_t, \tag{1}$$

with the ε_t's being uncorrelated random variables with zero mean and variance σ^2 (It should be understood that the parameters a, b and σ^2 refer to those at the end of the time series, and that they can change with time as more data become available). Then the weights $\{\theta_1, \ldots, \theta_M\}$ that minimize the expected squared revision,

$$R(\boldsymbol{\theta}) = \mathrm{E}\left(\sum_{i=1}^{M} \theta_i X_i - \sum_{i=1}^{N} w_i X_i\right)^2 ,$$

under the constraint $\sum_{i=1}^{M} \theta_i = 1$, are given by

$$\theta_i = w_i + \frac{1}{M}\sum_{j=M+1}^{2p+1} w_j + \frac{\left(i - \frac{M+1}{2}\right) D}{1 + \frac{M^3 - M}{12} D} \sum_{j=M+1}^{2p+1} \left(j - \frac{M+1}{2}\right) w_j ,$$

where $D = b^2/\sigma^2$. Note here that for M fixed, the index i varies from 1 to M.

In X-12-Arima, as in X-11, the order of the Henderson's moving average used to obtain an estimate of the trend-cycle is based on a "I/C-ratio" which is computed as

$$I/C = \frac{\Sigma\,|I_t - I_{t-1}|}{\Sigma\,|C_t - C_{t-1}|}$$

when the series X_t is expressed as the sum of a trend-cycle C_t and an irregular component I_t. Under the additional assumption that the ε_t's have a normal distribution, the value of D and I/C are related by the formula $D = 4/(\pi(I/C)^2)$ (see for example Doherty [2]). In X-12-Arima, $I/C = 1.0, 3.5$ and 4.5 are used with the 9, 13 and 23-term Henderson's averages respectively and for the Musgrave's surrogates.

Actually, the approach adopted by Musgrave can be easily generalized in different ways, especially by considering less specific models to describe the series and by choosing other kinds of benchmark symmetric weights w_i. Thus, suppose that the last $2p + 1$ data points of the time series follow a polynomial trend of degree d and that we have, for $t = -p, \ldots, +p$, $X_t = a_0 + a_1 t^1 + \ldots + a_d t^d + \varepsilon_t$, where $(\varepsilon_t)_t$ is a random stationary process with zero mean. With matricial notations, these assumptions can be rewritten as follows: $\mathbf{X} = \mathbf{A}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, with $\mathbf{X} = [X_{-p}, X_{-p+1}, \ldots, X_{+p}]'$, $\boldsymbol{\beta} = [a_0, a_1, \ldots, a_d]'$, $\boldsymbol{\varepsilon} = [\varepsilon_{-p}, \varepsilon_{-p+1}, \ldots, \varepsilon_{+p}]'$ and

$$\mathbf{A} = \begin{pmatrix} 1 & -p & (-p)^2 & \ldots & (-p)^d \\ 1 & -p+1 & (-p+1)^2 & \ldots & (-p+1)^d \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & +p & (+p)^2 & \ldots & (+p)^d \end{pmatrix}.$$

If $\mathbf{w} = (w_{-p}, \ldots, w_{+p})'$ denotes the vector of the benchmark symmetric moving average coefficients and $\boldsymbol{\theta} = (\theta_{-p}, \ldots, \theta_{+p})'$ the vector of the asymmetric moving average (with its coefficients θ_k equal to 0 if $p + 1 \le k \le f$), the revision criterion to be minimized can be then expressed as:

$$R(\boldsymbol{\theta}) = \mathrm{E}\left[\sum_{i=-p}^{+p} \theta_i X_i - \sum_{i=-p}^{+p} w_i X_i\right]^2 = (\boldsymbol{\theta} - \mathbf{w})'\mathbf{R}(\boldsymbol{\theta} - \mathbf{w})$$

where, as $\mathrm{E}(\boldsymbol{\varepsilon}) = 0$ and $\mathrm{Var}(\boldsymbol{\varepsilon}) = \boldsymbol{\Sigma}$, matrix $\mathbf{R}$ is given by :

$$\mathbf{R} = \mathrm{E}(\mathbf{X}\mathbf{X}') = \mathrm{E}\left[(\mathbf{A}\boldsymbol{\beta} + \boldsymbol{\varepsilon})(\mathbf{A}\boldsymbol{\beta} + \boldsymbol{\varepsilon})'\right] = \mathbf{A}\boldsymbol{\beta}\boldsymbol{\beta}'\mathbf{A}' + \boldsymbol{\Sigma}.$$

3.2 Towards the Generalized Direct Filter Approach

Heretofore, the way to derive moving averages regarding optimal revision purposes is guided by parametric modelings, which force to make more or less strong assumptions on the input time series. Conversely, a non-parametric approach can also be considered. By only assuming that the filter error, i.e. the difference between the benchmark symmetric filtered signal and the asymmetric filtered one, is a stationary process, the Direct Filter Approach (DFA, see [11, 12]) consists first in "translating" the revision criterion to be minimized in the frequency domain. For stationary processes X_t, the mean squared filter error can indeed be expressed as

$$\mathrm{E}\left(L_{\boldsymbol{\theta}} X_t - L_{\mathbf{w}} X_t\right)^2 = \int_0^{2\pi} \left| \widehat{L}_{\boldsymbol{\theta}}\left(e^{i\omega}\right) - \widehat{L}_{\mathbf{w}}\left(e^{i\omega}\right) \right|^2 dH_X\left(\omega\right), \tag{2}$$

where $H_X(\omega)$ is the unknown spectral density of X_t. The weights $\boldsymbol{\theta}$ of the asymmetric filter can then be computed by minimizing an estimation of the right-hand term of equation 2, based on a discretization of the interval $[0; 2\pi]$. The case of non-stationary integrated processes can be also handled within this framework, insofar as cointegration between both filtered signals can be easily satisfied by imposing linear constraints on vector $\boldsymbol{\theta}$, like those presented in section 2.

The main interest of this procedure lies in the fact that it is now possible to split the revision criterion into two distinct effects, one related to the gain and one related to the phase of the transfer functions. Indeed, as noticed by Wildi (see [13]), we have :

$$\begin{aligned}
&\left| \widehat{L}_{\boldsymbol{\theta}}\left(e^{i\omega}\right) - \widehat{L}_{\mathbf{w}}\left(e^{i\omega}\right) \right|^2 \\
&\qquad = \left(\rho_{\boldsymbol{\theta}}(\omega) - \rho_{\mathbf{w}}(\omega)\right)^2 + 2\rho_{\mathbf{w}}(\omega)\,\rho_{\boldsymbol{\theta}}(\omega)\left[1 - \cos\left(\varphi_{\boldsymbol{\theta}}(\omega)\right)\right] \\
&\qquad = \underbrace{\left(\rho_{\boldsymbol{\theta}}(\omega) - \rho_{\mathbf{w}}(\omega)\right)^2}_{\text{Gain effect}} + \underbrace{4\rho_{\mathbf{w}}(\omega)\,\rho_{\boldsymbol{\theta}}(\omega)\sin^2\left(\varphi_{\boldsymbol{\theta}}(\omega)/2\right)}_{\text{Phase shift effect}}.
\end{aligned} \tag{3}$$

The previous equations are derived from the fact that the phase function of a symmetric filter is real, so that its corresponding phase $\varphi_{\mathbf{w}}(\omega)$ is always equal to 0 or π. In our scope of interest, we naturally suppose the latter is actually null (i.e. the transfer function is positive), at least at low frequencies. Indeed, classical constraints imposed to the filters, like preserving constant trends, are in agreement with such an assumption, since, for instance, condition $\sum_k w_k = 1$ implies that $\widehat{L}_{\mathbf{w}}(0)$ be equal to 1 and, by continuity arguments, $\varphi_{\mathbf{w}}(\omega)$ be null at low frequencies.

Thanks to this factorization of the revision criterion, we are now able to deal with the trade-off between reliability and timeliness aspects, by emphasizing more or less the phase shift effect and by attaching varying importance to the different frequency components, depending on whether they appear in the spectrum of the initial time series or not. Thus, Wildi first suggests to modify the revision criterion in the following way,

$$\begin{aligned}
&\int_0^{2\pi} \left| \widehat{L}_{\boldsymbol{\theta}}\left(e^{i\omega}\right) - \widehat{L}_{\mathbf{w}}\left(e^{i\omega}\right) \right|^2 W(\omega)\, dH_X(\omega) \\
&\qquad + 4\lambda \int_0^{2\pi} \rho_{\mathbf{w}}(\omega)\,\rho_{\boldsymbol{\theta}}(\omega)\sin^2\left(\varphi_{\boldsymbol{\theta}}(\omega)/2\right) W(\omega)\, dH_X(\omega),
\end{aligned} \tag{4}$$

where λ is a tuning factor, balancing the gain and phase shift components, and $W(\omega)$ is a frequency weighting function. Increasing λ amounts to giving priority to the minimization of the phase error and so to the real-time detection of turning points - or more generally of trend changes.

As mentioned previously, the presence of phase shift effects is mainly due to the asymmetry of the filter, or equivalently to the fact that the tranfer function is complex. Consequently, another way to reduce them consists in giving an higher weight to the imaginary part of the transfer function in the criterion to be minimized. To generalize the DFA, Wildi has then proposed recently to replace $\widehat{L}_{\boldsymbol{\theta}}\left(e^{i\omega}\right)$ by $\Re\left(\widehat{L}_{\boldsymbol{\theta}}\left(e^{i\omega}\right)\right) + i\sqrt{1+4\lambda\widehat{L}_{\mathbf{w}}\left(e^{i\omega}\right)} \cdot \Im\left(\widehat{L}_{\boldsymbol{\theta}}\left(e^{i\omega}\right)\right)$ in (2), where $\Re$ and $\Im$ denote the real part and the imaginary part respectively. Developing expression (2) by accounting for this surrogate and weighting frequencies leads him eventually to minimize a discretized version of the following criterion:

$$\int_0^{2\pi} \left|\widehat{L}_{\boldsymbol{\theta}}\left(e^{i\omega}\right) - \widehat{L}_{\mathbf{w}}\left(e^{i\omega}\right)\right|^2 W(\omega)\, dH_X(\omega) + 4\lambda \int_0^{2\pi} \rho_{\mathbf{w}}(\omega)\, \rho_{\boldsymbol{\theta}}^2(\omega) \sin^2\left(\varphi_{\boldsymbol{\theta}}(\omega)\right) W(\omega)\, dH_X(\omega)\,. \quad (5)$$

Unlike the previous criterion given in (4), the latter is a quadratic function of the filter coefficients, making it easier to solve the minimization problem from a computational point of view.

4. A general unifying framework to derive linear filters

4.1 From the revision criteria to a very general optimization problem

The revision criterion on which Musgrave's approach and the DFA are based is nothing else but a distance measure between the output signal and what the latter would be if we were not confronted with the lack of observations in the future. Actually, the choice of the distance measure as well as the time series to which the output signal is compared can be widened. The ways to derive linear filters presented in section 3 are encompassed within the following more general optimization problem,

$$\begin{cases} \min\limits_{\boldsymbol{\theta}} \mathrm{E}\left[\left[\nabla^q\left(L_{\boldsymbol{\theta}}X_t - u_t\right)\right]^2\right] \\ \quad \text{subject to } \ \mathbf{C}\boldsymbol{\theta} = \mathbf{a}, \end{cases} \quad (6)$$

where ∇ is the discrete differentiation operator on time series, $\nabla X_t = X_t - X_{t-1}$, and, for any integer $q \geq 1$, $\nabla^q = \nabla \circ \nabla^{q-1} = \underbrace{\nabla \circ \ldots \circ \nabla}_{q \text{ factors}}$ (with ∇^0 denoting the identity function). The filter coefficients are chosen by minimizing the q-th difference of the discrepancy between the output signal and any benchmark time series $(u_t)_{t\in\mathbb{Z}}$. Moreover, they must satisfy some given linear constraints, for which the equality $\mathbf{C}\boldsymbol{\theta} = \mathbf{a}$ provides a matricial notation. The number of rows of matrix $\mathbf{C}$ and vector $\mathbf{a}$ is equal to the number of constraints. Because the latter are supposed to be linearly independent, we assume that $\mathbf{C}$ is a full row rank matrix.

The benchmark time series u_t is chosen to yield a rough estimate of the trend-cycle of the initial signal X_t. It can be either deterministic or stochastic and may depend itself on the filter coefficients. The order of differentiation and the linear constraints are chosen so that the filter satisfy expected properties like preserving trends and removing seasonalities. In this way, they can ensure that the process $\nabla^q\left(L_{\boldsymbol{\theta}}X_t - u_t\right)$ be stationary.

Moreover, beyond these considerations, the choice of modeling assumptions to apply to the initial time series is completely free. By varying these assumptions, we obtain a very large class of linear filters. Musgrave's moving averages and the DFA are obtained by choosing $q = 0$ and $u_t = L_{\mathbf{w}}X_t$. In the first case, a parametric model is used whereas, in the second case, the filter coefficients are computed thanks to a non-parametric estimation of the data spectral density. Actually, in the next subsection, we will see that more constraining assumptions can lead to well-known filters which depend neither on the data nor even on the date of estimation.

4.2 Classical particular cases

4.2.1 Fidelity

A first application of the general framework given in 4.1 provides simple moving averages suggested by Bongard (see [1]). Indeed, Bongard's approach consists in choosing moving averages having a strong attenuation effect on the irregular component, once trends have been preserved and seasonalities removed. In the decomposition of the unadjusted series, the residual is often modeled in the form of a white noise ε_t with zero expectation and constant variance σ^2. This white noise is transformed by the moving average into a sequence of random variables, ε_t^*, with a constant variance equal to $\sigma^2 \sum_{k=-p}^{+f} \theta_k^2$. Reducing the irregular component, and therefore its variance, then amounts to reducing the criterion $\sum_{k=-p}^{+f} \theta_k^2$. The output signal is supposed to be 'as close as possible' to the input signal where noise components are removed; that is why this criterion will be called here the 'Fidelity' criterion, denoted by $F(\boldsymbol{\theta})$. Considering additive white noises, the 'Fidelity' criterion is a very simple positive quadratic form,

$$F(\boldsymbol{\theta}) = \sum_{k=-p}^{+f} \theta_k^2 = \boldsymbol{\theta}'\mathbf{F}\boldsymbol{\theta},$$

with $\mathbf{F}$ being the identity matrix of order $p + f + 1$. This way of designing linear filters can be directly obtained by considering the general problem (6), with $q = 0$ and where u_t is the deterministic benchmark time series $u_t = \mathrm{E}(L_{\boldsymbol{\theta}} X_t)$.

4.2.2 Smoothness

X-12-ARIMA uses Henderson's moving averages to extract the trend-cycle from an estimate of the seasonally adjusted series (Tables B7, C7, D7, D12). In the additive case, the model which is considered then becomes $X_t = TC_t + I_t$. Henderson suggests to use a moving average derived from a criterion which ensures a smooth estimation of the trend-cycle. Let us consider the Dirac time series $\delta_t^{t_0}$, equal to 1 at instant $t = t_0$ and to 0 at any other instant. Its transform by a moving average $L_{\boldsymbol{\theta}}$ of order $p + f + 1$ with coefficients $\{\theta_k\}$, is given by:

$$L_{\boldsymbol{\theta}} \delta_t^{t_0} = \begin{cases} \theta_{t_0 - t} & \text{if } t_0 - f \leq t \leq t_0 + p \\ 0 & \text{otherwise} \end{cases} .$$

This transform will therefore be smooth if the coefficient curve of the moving average is not too irregular. Noting that the vector space of time series is generated by the Dirac ones, since any time series X_t can be written as $X_t = \sum_{t_0 \in \mathbb{Z}} X_{t_0} \delta_t^{t_0}$, Henderson [5, 6] proposes to use the quantity $S = \sum_k \left(\nabla^3 \theta_k\right)^2$ to measure the 'flexibility' of the coefficient curve. The notation ∇, introduced in section 4.1 as a linear operator on time series, is extended with a similar meaning onto the set of vector coefficients $\boldsymbol{\theta}$, i.e. $(\nabla \boldsymbol{\theta})_k = \nabla \theta_k = \theta_k - \theta_{k-1}$. Quantity S vanishes when the coefficients $\{\theta_k\}$ are located along a parabola. In the general case, it measures the difference between a parabola and the form of the coefficient curve. Henderson then looked for centered averages that preserve quadratic polynomials and minimize S.

Minimizing Henderson's criterion, or more generally the 'Smoothness criterion'

$$S(\boldsymbol{\theta}) = \sum_k \left(\nabla^q \theta_k\right)^2,$$

aims at getting quite regular output signals. Notice that $S(\boldsymbol{\theta})$, as the previous 'Fidelity' criterion, is also a positive quadratic form over $\mathbb{R}^{p+f+1}$. For instance, if we consider Henderson's case $q = 3$, we have $S(\boldsymbol{\theta}) = \boldsymbol{\theta}'\mathbf{S}\boldsymbol{\theta}$, where $\mathbf{S}$ is the following symmetric Toeplitz matrix :

$$\mathbf{S} = \begin{pmatrix} 20 & -15 & 6 & -1 & 0 & \dots & 0 \\ -15 & 20 & -15 & 6 & \ddots & \ddots & \vdots \\ 6 & -15 & 20 & \ddots & \ddots & \ddots & 0 \\ -1 & 6 & \ddots & \ddots & \ddots & 6 & -1 \\ 0 & \ddots & \ddots & \ddots & 20 & -15 & 6 \\ \vdots & \ddots & \ddots & 6 & -15 & 20 & -15 \\ 0 & \dots & 0 & -1 & 6 & -15 & 20 \end{pmatrix}.$$

Actually, let us extend the notation of the filter coefficients by stating that $\theta_k = 0$ for any integer $k \notin [-p; +f]$. We have then : $\nabla^q (L_{\boldsymbol{\theta}} X_t) = (-1)^q \sum_{k \in \mathbb{Z}} (\nabla^q \theta_k) X_{t+k-q}$. Thus, if we consider similar assumptions on the irregular component to those made for the Fidelity criterion in section 4.2.1, Henderson's approach then appears to be a particular case of problem (6) where u_t is the deterministic benchmark time series $u_t = \mathrm{E}(L_{\boldsymbol{\theta}} X_t)$ again, but q is now a positive integer.

Other classical linear filters, like the Hodrick-Prescott filter, can also be considered as particular cases of problem (6).

5. The Timeliness criterion

5.1 Minimizing the phase shift

As suggested in section 2.3, an approach to reduce the observed phase shift between any input signal and its corresponding output signal could consist in searching for moving averages which introduce a small phase shift when they are applied to harmonic signals $\left(e^{i\omega t}\right)_{t \in \mathbb{Z}}$. In other words, we are going to search for vectors $\boldsymbol{\theta}$ such that the phase $\varphi_{\boldsymbol{\theta}}(\omega)$ of filter $L_{\boldsymbol{\theta}}$ be close to 0. Such an approach should provide moving averages with *a priori* satisfying properties in terms of phase shift, whatever the input signal to which they are applied.

If a vector $\boldsymbol{\theta}$ is such that $\varphi_{\boldsymbol{\theta}}(\omega_0) = 0$, then it guarantees that both signals $\left(e^{i\omega_0 t}\right)_{t \in \mathbb{Z}}$ and $\left(L_{\boldsymbol{\theta}} e^{i\omega_0 t}\right)_{t \in \mathbb{Z}}$ be in phase, but this property is no more satisfied for any pair of signals $\left(e^{i\omega t}\right)_{t \in \mathbb{Z}}$ and $\left(L_{\boldsymbol{\theta}} e^{i\omega t}\right)_{t \in \mathbb{Z}}$ as soon as we consider a frequency ω different from ω_0. In fact, as the transfer function is regular and the phase is supposed to be wihin the range $]-\pi; \pi]$, it can be only stated that the phase $\varphi_{\boldsymbol{\theta}}(\omega)$ is also a regular function and thus is close to 0 in a neighbourhood of ω_0. As any stationary signal can be decomposed into linear combinations of harmonic signals characterized by different frequencies, the filter coefficients $\boldsymbol{\theta}$ must be actually chosen among those for which the phase function is close to 0, whatever the frequency ω. Practically, we will rather consider the range of frequencies $]\omega_1; \omega_2[$ which belong to the spectrum of stationary components of time series to which the moving average $L_{\boldsymbol{\theta}}$ will be applied, with $0 \leq \omega_1 \leq \omega_2 \leq 2\pi$.

Besides, when filtering, the harmonic component of the input signal at frequency ω is multiplied by an amplitude factor equal to the modulus function. Therefore, it is natural to consider that the impact of the phase shift on the harmonic component at frequency ω will be all the more significant so as the value $\rho_{\boldsymbol{\theta}}(\omega)$ taken by the modulus function is high.

As a consequence, in order to reduce the phase shift effects between input and output signals, we suggest the use of a moving average whose coefficients minimize the following criterion,

$$\int_{\omega_1}^{\omega_2} f\left[\rho_{\boldsymbol{\theta}}(\omega); \varphi_{\boldsymbol{\theta}}(\omega)\right] d\omega, \qquad (7)$$

where the function f, defined over $[\,0; +\infty\,[\,\times\,]-\pi; \pi\,]$ is chosen beforehand and satisfies the six following conditions:

1. $f \geq 0$
2. $f(\rho, 0) = 0$
3. $f(0, \varphi) = 0$
4. $f(\rho, \varphi) = f(\rho, -\varphi)$
5. $\frac{\partial f}{\partial \rho} \geq 0$
6. $\varphi \cdot \frac{\partial f}{\partial \varphi} \geq 0$

From now on, the function f is called the penalty function. Conditions 1 and 2 assure that this function is minimized when the phase is null. Conditions 4 and 6 mean that, in the criterion (7), the phase is penalized only in function of its distance from 0, all the more so as such a distance is high. Moreover, condition 5 means that, the more a frequency is amplified by the filter, the more the corresponding phase shift should be reduced. Eventually, in the limit case of a frequency cut off by the filter, condition 3 stipulates that it is unnecessary to give a non null weight to the corresponding phase in the criterion.

5.2 The choice of a convenient penalty function

Now, the penalty function has to be chosen. It would be desirable to find a function satisfying the six previous conditions such that the problem consisting in minimizing (7) can be solved analytically or, if it is not possible, by using numerical algorithms. In the latter case, we must then be able to prove the existence of both a theoretical solution and a numerical algorithm converging quickly to this solution. A penalty function f providing a convex criterion would be suitable since minimizing the quantity given in (7) would then become a classical convex optimization problem for solving of which some converging numerical algorithms are well-known (gradient methods,...).

Proposition 1 (A first family) *: For any $k > 0$ and $\ell > 0$, the function $(\rho, \varphi) \longmapsto \rho^k |\varphi|^\ell$ satisfies the six conditions imposed on the penalty function appearing in criterion (7).*

However, it seems better to force the second parameter ℓ to be greater than or equal to 1. Otherwise, the partial derivative $|\partial f / \partial \varphi|$ would tend to $+\infty$ when φ tends to 0 ; non-null phases close to 0 would be too much penalized.

Proposition 2 (A second family) *: For any $k > 0$ and $\ell > 0$, the function $(\rho, \varphi) \longmapsto \rho^k \left|\sin\left(\frac{\varphi}{2}\right)\right|^\ell$ satisfies the six conditions imposed on the penalty function appearing in criterion (7).*

Once again, we can recommend the use of a function with a parameter ℓ greater than or equal to 1. Let us notice that functions of this second family are $2\pi-$periodic with respect to the variable φ (so that it would be no more necessary to force the phase to belong to $]-\pi;\pi\,]$).

Here, we suggest choosing the penalty function given in proposition 2 with parameters $k=1$ et $\ell=2$. In other words,

$$f\left[\rho_{\boldsymbol{\theta}}(\omega);\varphi_{\boldsymbol{\theta}}(\omega)\right]=\rho_{\boldsymbol{\theta}}(\omega)\sin^2\left(\frac{\varphi_{\boldsymbol{\theta}}(\omega)}{2}\right)=\frac{1}{2}\left[\rho_{\boldsymbol{\theta}}(\omega)-\Re\left(\widehat{L}_{\boldsymbol{\theta}}\left(e^{i\omega}\right)\right)\right] \tag{8}$$

As shown later, such a penalty function is worthwhile inasmuch as it makes the minimization of criterion (7) solvable, at least with classical numerical algorithms. But notice also that it corresponds exactly to the phase shift effect appearing in (3), which is then emphasized in the revision criterion (4) for the Direct Filter Approach. However, in this section, the filter is supposed to be independent from the data and its phase shift properties are analysed directly and not compared to those of a convenient symmetric filter.

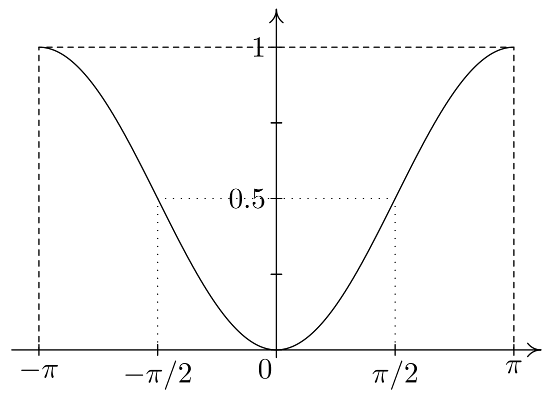

Figure 2: Penalty function $\varphi\longmapsto f(1,\varphi)=\sin^2\left(\frac{\varphi}{2}\right)$

Let now $\mathbf{Co}(\omega)$ and $\mathbf{Si}(\omega)$ denote the vectors defined by $\left[\cos(-p\omega),\ldots,\cos(f\omega)\right]'$ and $\left[\sin(-p\omega),\ldots,\sin(f\omega)\right]'$ respectively. Let also $\boldsymbol{\Omega}(\omega)$ be the square matrix of order $p+f+1$ whose entry lying in the $k-$th row and the $\ell-$th column is given by $\boldsymbol{\Omega}_{k\ell}=\cos\left[(k-\ell)\,\omega\right]$ for any indices $-p\leq k,\ell\leq +f$. Note that we have then $\Re\left(\widehat{L}_{\boldsymbol{\theta}}\left(e^{i\omega}\right)\right)=\mathbf{Co}(\omega)'\boldsymbol{\theta}$, $\Im\left(\widehat{L}_{\boldsymbol{\theta}}\left(e^{i\omega}\right)\right)=\mathbf{Si}(\omega)'\boldsymbol{\theta}$ and $\boldsymbol{\Omega}(\omega)=\mathbf{Co}(\omega)\,\mathbf{Co}(\omega)'+\mathbf{Si}(\omega)\,\mathbf{Si}(\omega)'$, so that $\rho^2_{\boldsymbol{\theta}}(\omega)=\boldsymbol{\theta}'\boldsymbol{\Omega}(\omega)\,\boldsymbol{\theta}$. Thus, matrix $\boldsymbol{\Omega}(\omega)$ is symmetric, positive semi-definite but non-invertible (its rank is lesser than or equal to 2).

Considering the penalty function given in (8), criterion (7) can be rewritten in the following way:

$$T_0(\boldsymbol{\theta})=\int_{\omega_1}^{\omega_2}\sqrt{\boldsymbol{\theta}'\boldsymbol{\Omega}(\omega)\,\boldsymbol{\theta}}\,d\omega-\left(\int_{\omega_1}^{\omega_2}\mathbf{Co}(\omega)\,d\omega\right)'\boldsymbol{\theta} \tag{9}$$

Proposition 3 *Criterion $T_0(\boldsymbol{\theta})$ given in (9) is a convex function of the vector of coefficients $\boldsymbol{\theta}$.*

The objective mentioned at the beginning of this section is achieved. Proposition 3 implies that the phase shift minimization problem for the penalty function given in (8), as a convex optimization problem, has a solution and is solvable by using for instance a gradient method. For this purpose, let us also give the expression of the gradient of criterion $T_0(\boldsymbol{\theta})$.

Proposition 4 *The gradient of criterion $T_0(\boldsymbol{\theta})$ given in (9) is given by:*

$$\frac{\partial T_0(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \left(\frac{1}{2} \int_{\omega_1}^{\omega_2} \frac{\boldsymbol{\Omega}(\omega)\boldsymbol{\theta}}{\sqrt{\boldsymbol{\theta}'\boldsymbol{\Omega}(\omega)\boldsymbol{\theta}}} \, d\omega \right) - \left(\int_{\omega_1}^{\omega_2} \mathbf{Co}(\omega) \, d\omega \right)$$

5.3 Variation on the Timeliness criterion

In this section, we propose releasing some constraints on the penalty function to get a more practical Timeliness criterion from a computational point of view. As already said in section 3.2, an other way to restrict the phase shift effects consists in reducing the imaginary part of the transfer function. By doing so, this boils down to considering the phase function *modulo* π and consequently to tolerate antiphases between input and output signals. Geometrically, this merely means that the distance from the transfer function to the real line in the complex plane is reduced as much as possible, whereas mathematically the penalty function f must now satisfy the six conditions given in section 5.1, but with a slight restriction. Indeed, the sixth condition has to be valid only for phases φ in the interval $[-\pi/2; \pi/2]$ and is completed by a seventh one, stipulating that for any phase $\varphi \in \,]\,\pi/2; \pi\,]$, $f(\rho, \varphi) = f(\rho, \pi - \varphi)$. A family of penalty functions satisfying the seven conditions is given below:

Proposition 5 *For any $k > 0$ and $\ell > 0$, the function $(\rho, \varphi) \longmapsto \rho^k |\sin(\varphi)|^\ell$ satisfies the seven conditions imposed on the penalty function appearing in the problem where phase shift effects are minimized modulo π.*

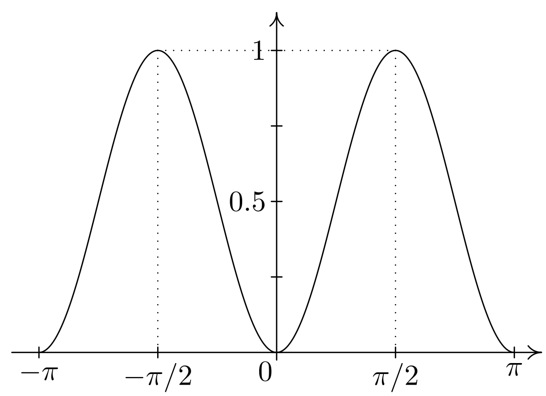

Figure 3: Penalty function $\varphi \longmapsto f(1, \varphi) = \sin^2(\varphi)$

Among these functions, the following one, obtained for $k = \ell = 2$,

$$f[\rho_{\boldsymbol{\theta}}(\omega); \varphi_{\boldsymbol{\theta}}(\omega)] = \rho_{\boldsymbol{\theta}}^2(\omega) \sin^2(\varphi_{\boldsymbol{\theta}}(\omega)) = \left[\Im\left(\widehat{L}_{\boldsymbol{\theta}}\left(e^{i\omega}\right)\right)\right]^2, \tag{10}$$

is particularly worthwhile because it makes the minimization of the Timeliness criterion analytically solvable. Indeed, the latter, denoted by $T(\boldsymbol{\theta})$, is simply a quadratic form :

$$T(\boldsymbol{\theta}) = \int_{\omega_1}^{\omega_2} y_{\boldsymbol{\theta}}^2(\omega) \, d\omega = \boldsymbol{\theta}'\mathbf{T}\boldsymbol{\theta}, \tag{11}$$

where $\mathbf{T} = \int_{\omega_1}^{\omega_2} \mathbf{Si}(\omega)\,\mathbf{Si}(\omega)'\, d\omega$ is a square matrix of order $p+f+1$. After a few calculations, it can be shown that its entries are given for $-p \leq k, \ell \leq f$ by:

$$\mathbf{T}_{kl} = \begin{cases} \frac{\sin((k-\ell)\omega_2) - \sin((k-\ell)\omega_1)}{2(k-\ell)} - \frac{\sin((k+\ell)\omega_2) - \sin((k+\ell)\omega_1)}{2(k+\ell)} & \text{if } |k| \neq |\ell| \text{ and } k\ell \neq 0 \\ \frac{\omega_2-\omega_1}{2} - \frac{\sin(2k\omega_2) - \sin(2k\omega_1)}{4k} & \text{if } k = \ell \text{ and } k\ell \neq 0 \\ \frac{\sin(2k\omega_2) - \sin(2k\omega_1)}{4k} - \frac{\omega_2-\omega_1}{2} & \text{if } k = -\ell \text{ and } k\ell \neq 0 \\ 0 & \text{if } k\ell = 0 \end{cases}$$

Not surprisingly, controling the phase shift effects by the means of the imaginary part of the transfer function has led us to a criterion (10) which is exactly the non-data-driven equivalent of the Generalized Direct Filter Approach summed up in equation (5). Its main drawback lies in the lack of penalization of antiphases. However, this is not really knotty. Indeed, for the same reasons as those invoked in section 3.2, the transfer function is close to 1 and therefore the phase function close to 0 at low frequencies. Thus, antiphases can only occur at higher frequencies, but the constraints usually imposed to the filter coefficients, regarding trends or seasonalities, aim precisely at removing or at least strongly reducing the high frequency components of the input signals.

6. Operational procedure and simulations

6.1 A mixed criterion

It is important to notice that there are always an endless number of solutions $\boldsymbol{\theta}$ to the minimization of the Timeliness criterion. For instance, criterion T_0 is convex but not strictly convex, whereas criterion T is only a *semi-definite* positive quadratic form. For instance, the latter is null and so reaches its minimum exactly on the whole vector subspace of symmetric vectors, i.e. of vectors $\boldsymbol{\theta}$ such that $\theta_k = \theta_{-k}$ for any $|k| \leq \min(p, f)$ and $\theta_k = 0$ otherwise.

Classical procedures to choose the coefficients of a moving average have been based on the minimization of the Fidelity and/or the Smoothness criteria. A quite natural idea could consist now in mixing them with the Timeliness criterion to account for the various biases induced by a moving average. Using the notations introduced in section 4.1, 4.2 and 5.3 to depict the linear constraints regarding trends and seasonalities and to compute the previously mentioned criteria, we suggest the following operational procedure to derive moving averages minimizing phase shift effects:

Problem 1 (The operational FST procedure)

$$\begin{cases} \min_{\boldsymbol{\theta}} J(\boldsymbol{\theta}) \\ \textit{subject to } \ \mathbf{C}\boldsymbol{\theta} = \mathbf{a} \end{cases}, \ \textit{where } J(\boldsymbol{\theta}) = \alpha \cdot F(\boldsymbol{\theta}) + \beta \cdot S(\boldsymbol{\theta}) + \gamma \cdot T(\boldsymbol{\theta}),$$

the 3 real numbers α, β, γ being given in the interval $[0; 1]$ and satisfying two conditions, $\alpha + \beta + \gamma = 1$ and $\alpha\beta \neq 0$.
Such an optimization problem has a unique solution.

The first condition on real numbers α, β, γ means that the mixed criterion $J(\boldsymbol{\theta})$ is a convex linear combination of the Fidelity, Smoothness and Timeliness ones. The second condition guarantees that $J(\boldsymbol{\theta})$ is a strictly convex function, that is why the solution to problem 1 is unique. Incidentally, the latter is given by $\boldsymbol{\theta} = \mathbf{J}^{-1}\mathbf{C}'(\mathbf{C}\mathbf{J}^{-1}\mathbf{C}')^{-1}\mathbf{a}$, where $\mathbf{J} = \mathbf{F} + \mathbf{S} + \mathbf{T}$. Obviously, Timeliness criterion $T(\boldsymbol{\theta})$ could be replaced by the more accurate one $T_0(\boldsymbol{\theta})$ presented in formula (9), but resolution of problem 1 would be then no more analytical.

The problems of the optimal choice of the coefficients α, β, and γ as well as the design of the matrix of linear constraints will be the core of a next paper.

6.2 An application to real data

Below is an application to a time series based on real data. Both the raw series (black line) and the series smoothed by the classical asymmetric Henderson moving average with $p = 9$ et $f = 3$ (purple line) are represented in figure 4. On figure 5, we can see the same raw time series (black line) and the series now smoothed by the 9-3 Henderson moving average minimizing phase-shift (purple line). In the areas circled in purple, the moving average minimizing phase-shift yields a significantly better adjustment in terms of 'fidelity' and 'timeliness', compared to usual Henderson filters applied in seasonal adjustment softwares.

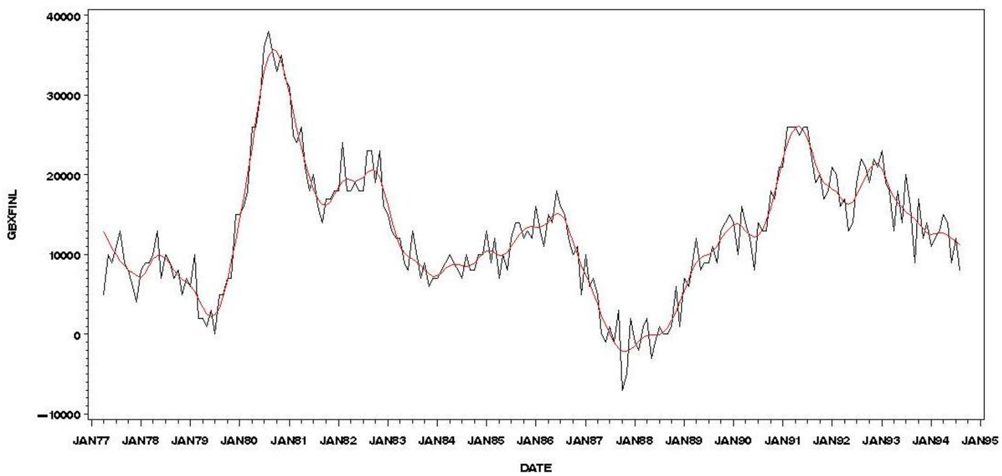

Figure 4: Raw time series (black) and time series smoothed by classical 9-3 Henderson moving average (purple)

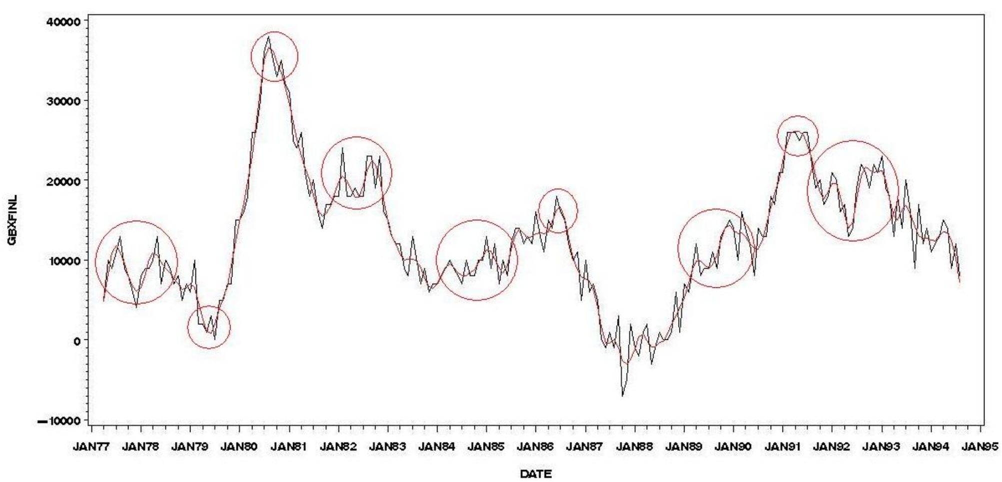

Figure 5: Raw time series (black) and time series smoothed by the 9-3 Henderson moving average minimizing phase-shift (purple)

Acknowledgments

This work was part of a research program supported by Eurostat, the statistical office of the European Community. The authors would like to thank Marc Wildi, from ZHAW

(Zürcher Hochschule für Angewandte Wissenschaften), for his kind help and his very useful comments.

References

[1] Bongard, J. (1962), "Quelques remarques sur les moyennes mobiles," *L'limination des variations saisonnires l'aide de calculatrices lectroniques*, OCDE, 1962, 389–427, Paris.

[2] Doherty, M. (2001), "Surrogate Henderson Filters in X-11," *Australian and New Zealand Journal of Statistics*, 43, 385–392. (Initially published as a working paper in 1992).

[3] Gray, A., Thomson, P. (2002), "On a Family of Moving Average Trend Filters for the Ends of Series," *Journal of Forecasting*, 21, 125–149. (Initially published as a working paper in 1996.)

[4] Grun-Rehomme, M., Ladiray, D. (1994), "Moyennes mobiles centrées et non centrées: construction et comparaison," *Revue de Statistique Appliquée*, vol XLII, 3, 33–61, Paris.

[5] Henderson, R. (1916), "Note on Graduation by Adjusted Average," *Transactions of the Actuarial Society of America*, 17, 43–48.

[6] Henderson, R. (1924), "A New Method of Graduation," *Transactions of the Actuarial Society of America*, 25, 29–40.

[7] Ladiray, D., Quenneville, B. (2001), "*Understanding X-11*," Springer-Verlag, Lecture Notes in Statistics 158, New York.

[8] Musgrave, J.C. (1964), "A Set of End Weights to End All End Weights," *Unpublished internal note*, Washington: US Bureau of the Census.

[9] Quenneville, B., Ladiray, D. (2000), "Locally Adaptive Trend-Cycle Estimation for X-11," *Proceedings of the International Conference on Establishment Surveys II*, Buffalo, NY, June 2000.

[10] Quenneville, B., Ladiray, D., Le Franois, B. (2003), "A Note on Musgrave Asymmetrical Trend-Cycle Filters," *International Journal of Forecasting*, 19, 727–734.

[11] Wildi, M. (1998), "Detection of Compatible Turning Points and Signal Extraction for Non-Stationary Time Series," *Operation Research Proceedings*, Springer.

[12] Wildi, M. (2011), "Signal Extraction: Efficient Estimation, 'Unit-Root' Tests and Early Detection of Turning Points," Springer-Verlag, Lecture Notes in Economic and Mathematical Systems, 547, Berlin Heidelberg.

[13] Wildi, M. (2011), "Real Time Trend Extraction and Seasonal Adjustment: a Generalized Direct Filter Approach," *To be published*.