# Reducing revisions in real time trend-cycle estimation

Estela Bee Dagum* Silvia Bianconcini†

**Abstract**
Recently, reproducing kernel Hilbert spaces have been introduced to provide a common approach for studying several nonparametric estimators used for smoothing time series data (Dagum and Bianconcini, 2008 and 2011). Based on this methodology, Bianconcini and Quenneville (2010) focused on the properties of the Henderson reproducing kernels when the filters are adapted at the end of the sample period, and with particular emphasis on the influence of the kernel order and bandwidth parameter. In this paper, we design a family of trend filters applied for real time estimation that are optimal in terms of reducing revisions when new observations are added to the series, and that are characterized by a fast detection of true turning points.

**Key Words:** End-point trend-cycle estimation, Henderson kernels, revision minimization, turning point detection.

## 1. Introduction

Real time trend-cycle estimation is of outmost importance for policy makers at any level of economic activity because it enables to assess the stage of the business cycle at which the economy stands. Hence, current trend-cycle estimates subject to frequent and high revisions are disliked by policy makers, particularly, if the revisions indicate a change of direction of the cyclical movement.

The majority of the seasonal adjustment software used by government agencies and statistical bureau estimate the trend-cycle component using moving average techniques. In particular, the linear smoother developed by Henderson (1916) has a long tradition for trend-cycle estimation in economic time series. Henderson filters are still employed for trend estimation in the X-11 cascade filter, and as such are an integral part of the X-12-ARIMA procedure, the seasonal adjustment method officially adopted in the US, Canada, United Kingdom, and many other countries. The importance of real time trend-cycle estimation has been stressed in several recent studies that focus on the properties of the asymmetric Henderson filters applied to the most recent observations (see, among others, Doherty, 2001; Quenneville et al., 2003; Dagum and Bianconcini, 2008 and 2011; Bianconcini and Quenneville, 2010). A problem that is common to all the methodologies is the extent to which these trend estimates are accurate and precise. At the end of the sample period, the data are subject to revisions when additional observations are added because of the use of asymmetric filters, hence these filters have to be optimal in some way.

As a local polynomial estimator, the Henderson asymmetric filters can be derived through automatic adaptations at the end of the sample period. This implies that the bias and the variance near the boundaries of the series are of the same order of magnitude as in the interior (see *e.g.* Fan and Gijbels, 1996). It turns out that the variance inflation resulting from the one-sided real time Henderson filter is very high, and that the filter is strongly localized at the current observation. In order to overcome such limitations, alternative strategies have been applied in the literature.

The idea embodied in the X11/X12ARIMA seasonal adjustment procedure is to apply the

*Department of Statistics, University of Bologna, via delle Belle Arti 41, 40126 Bologna, Italy
†Department of Statistics, University of Bologna, via delle Belle Arti 41, 40126 Bologna, Italy

symmetric two-sided Henderson filter to the series extended by forecasts (Dagum, 1982). This strategy is safer, provided that we are capable of producing optimal forecasts from according to some time series model of the ARIMA class. An intuitive and easily proven fact is that if the forecasts are optimal in the mean square error sense, then the variance of the revisions is minimum (Wallis, 1983).

In X11 based seasonal adjustment procedures, the end-point estimates of the trend-cycle are obtained using asymmetric filters that were developed by Musgrave (1964), specifically to minimize revisions for a certain class of time series. Doherty (2001) is the most complete study on how Musgrave surrogates, *i.e.* the X11 asymmetric trend-cycle moving averages, are calculated. Quenneville and Ladiray (2000) also studied the dependence of these surrogates on a parameter that is restricted to predetermined values in X11. They found that the differences in terms of revisions between the Musgrave averages with the estimated parameter and the Henderson average with forecasts from a simple ARIMA model can be very small.

Even if Musgrave asymmetric filters are optimal in terms of minimum mean square revision of the estimates, they are derived following a different optimization criteria with respect to the symmetric Henderson filter, with the consequence that the asymmetric filters do not converge monotonically to the symmetric one. At this regard, Dagum and Bianconcini (2008 and 2011) provided a different characterization of the Henderson weights within the Reproducing Kernel Hilbert Space (RKHS) methodology. According to this approach, a continuous kernel representation of the Henderson filter is obtained, and the same kernel function is used to derive both symmetric and asymmetric weights. The density function (*i.e.* a second order kernel) embedded on the linear filter is firstly determined. This is the starting point for obtaining higher order kernels, which are based on the product of the density function and its orthonormal polynomials. Bianconcini and Quenneville (2010) showed that, for each kernel order, the asymmetric filters can be derived coherently with the corresponding symmetric weights. In the particular case of the currently applied asymmetric Henderson filters, those obtained by means of the RKHS are shown to have superior properties from the viewpoint of signal passing, noise suppression, and size of revisions.

The main purpose of this paper is real time estimation of the underlying trend in a time series by means of asymmetric Henderson weights derived using the RKHS methodology. In particular, a family of end filters is constructed using a minimum revision criterion. We analyze the statistical properties of these asymmetric filters in terms of both size of revisions and time to detect turning points, paying particular attention on the selection of the kernel function and bandwidth parameter.

The paper is structured as follows. Section 2 discusses the reproducing kernel representations of the classical Henderson filter. Section 3 illustrates the properties of boundary Henderson kernels, obtained by applying the so called cut-and-normalize method. We also derive a family of optimal boundary Henderson kernels, optimal in the sense of minimizing the size of revisions in the final estimates, sharing also the property of fast turning point detection. In particular, we address inferential issues related to the optimal choice of the kernel function and of the bandwidth parameter. Finally, Section 5 provides an empirical application.

## 2. Reproducing Henderson kernels for trend-cycle estimation

Let $\{y_t, t = 1, 2, ..., N\}$ denote the input series, that is assumed to be decomposed into the sum of a systematic component, called the signal (or nonstationary mean) $g_t$, plus an erratic component $u_t$, called the noise, such that

$$y_t = g_t + u_t. \quad (1)$$

The noise $u_t$ is assumed to be either a white noise, $WN(0, \sigma_u^2)$, or, more generally, to follow a stationary and invertible AutoRegressive Moving Average (ARMA) process.
If $\{y_t, t = 1, 2, ..., N\}$ is seasonally adjusted or without seasonality, the signal $g_t$ represents the trend and cyclical components, usually referred to as trend-cycle for they are estimated jointly. The trend-cycle can be deterministic or stochastic, and can have a global or a local representation. It can be represented *locally* by a polynomial of degree $p$ of a variable $j$, which measures the distance between $y_t$ and the neighboring observations $y_{t+j}$. Equivalently, the trend-cycle estimate $\hat{g}_t$ is a weighted average applied in a moving manner, such that

$$\hat{g}_t = \sum_{j=-m}^{m} w_j y_{t+j} \tag{2}$$

where $w_j$, $j < N$, denotes the weights to be applied to the observations $y_{t+j}$ to get the estimate $\hat{g}_t$ for each point in time $t = 1, 2, ..., N$. Several nonparametric estimators, based on different sets of weights $w_j, j = -m, ...m$ have been developed in the literature.

Henderson's starting point is the requirement that the filter should reproduce a cubic polynomial trend without distortion. He proved that three alternative smoothing criteria give the same formula, as shown explicitly by Kenny and Durbin (1982) and Gray and Thomson (1996). These criteria are: (1) minimization of the variance of the third differences of the series smoothed by the application of the moving average ($\min var(\Delta^3 \hat{y}_t)$); (2) minimization of the sum of squares of the third differences of the coefficients of the moving average formula ($\min \sum_{j=-m}^{m} (\Delta^3 w_j)^2$); (3) fitting a cubic polynomial by weighted least squares to the observations $y_{t+j}, j = -m, ..., m$, where the weights are chosen as to minimize the sum of squares of their third differences. Representing the latter by $W_j, j = -m, ..., m$, where $W_j = W_{-j}$, the problem is the minimization of

$$\sum_{j=-m}^{m} W_j [y_{t+j} - a_0 - a_1 j - a_2 j^2 - a_3 j^3]^2, \tag{3}$$

where the solution for the constant term $\hat{a}_0$ is the smoothed observation $\hat{g}_t$. Henderson (1916) showed that $\hat{g}_t$ is given by

$$\hat{g}_t = \sum_{j=-m}^{m} \phi(j) W_j y_{t+j}, \tag{4}$$

where $\phi(j)$ is a cubic polynomial whose coefficients have the property that the linear filter reproduces the data if they locally follow a cubic. Henderson also proved the converse: if the coefficients of a cubic-reproducing summation formula $\{w_j, j = -m, ..., m\}$ do not change their sign more than three times within the filter span, then the formula can be represented as a local cubic smoother with weights $W_j > 0$ and a cubic polynomial $\phi(j)$, such that $\phi(j) W_j = w_j$.
Henderson (1916) measured the amount of smoothing of the input series by $\sum (\Delta^3 y_t)^2$ or equivalently by the sum of squares of the third differences of the weight diagram, $\sum (\Delta^3 w_j)^2$. The solution is that resulting from the minimization of a cubic polynomial function by weighted least squares with

$$W_j \propto \{(m+1)^2 - j^2\}\{(m+2)^2 - j^2\}\{(m+3)^2 - j^2\} \tag{5}$$

as the weighting penalty function of criterion (3) above. This optimality result has been rediscovered several times in modern literature, usually for asymptotic kernel variants. At this regard, Loader (1999) showed that the Henderson weights represent a finite sample variant of a kernel with second order vanishing moments which minimizes the third derivative of the function given by Muller (1984). In particular, Loader showed that for large $m$, the weights of Henderson's ideal penalty function $W_j$ are approximately $m^6 W(j/m)$, where $W(j/m)$ is the triweight density function. He concluded that, for very large $m$, the weight diagram is approximately $(315/512) * W(j/m)(3 - 11(j/m)^2)$ equivalent to the kernel given by Muller (1984).

Dagum and Bianconcini (2008) provided different kernel characterizations of the Henderson filter following the Reproducing Kernel Hilbert Space (RKHS) methodology.
A RKHS is a Hilbert space characterized by a kernel that reproduces, via an inner product, every function of the space or, equivalently, a Hilbert space of real-valued functions with the property that every point evaluation functional is bounded and linear. Berlinet and Thomas-Agnan (2003) showed how the space of polynomials of degree at most $p$ is a RKHS. At this regard, the minimization problem (3) can be rewritten in continuous time as follows

$$\min_{g \in \mathbf{P}_3} \|y - g\|^2_{\mathbf{P}_3} = \int_T (y(t-s) - g(t-s))^2 f_0(s) ds \tag{6}$$

where $\| \cdot \|^2_{\mathbf{P}_3}$ denotes the norm in the space of polynomials of degree at most 3, $\mathbf{P}_3$, and $f_0$ is the density corresponding to the weighting function $W_j$.

***Theorem 1.*** *The minimization problem (6) has a unique and explicit solution given by*

$$\hat{g}(t) = \int_T y(t-s) K_4(s) ds \tag{7}$$

*where $K_4$ is the fourth (also third for the symmetry of $f_0$) order kernel representation of the Henderson filter.*

The reader is referred to Dagum and Bianconcini (2008) for a detailed proof of Theorem 1. The kernel $K_4$ can be written as products of the reproducing kernel $R_3(t, .)$ of the space $\mathbf{P}_3$ and the density function $f_0$, that has finite moments up to order 8. In particular, using the Christoffel-Darboux formula, for the sequence $(P_i)_{0 \leq i \leq 3}$ of orthonormal polynomials with respect to the density $f_0$,

$$K_4(t) = \sum_{i=0}^{3} P_i(t) P_i(0) f_0(t). \tag{8}$$

The density corresponding to $W_j$ and its orthonormal polynomials have to be determined. As shown by Dagum and Bianconcini (2008), the former is given by

$$f_{0H}(t) = \frac{(m+1)}{k} W((m+1)t), \quad t \in [-1, 1] \tag{9}$$

where $k = \int_{-m-1}^{m+1} W(j) dj$, and $j = (m+1)t$. On the other hand, the polynomials were computed by solving the Hankel system based on the moments of the density $f_{0H}$ (Brezinski 1980).
The kernel $K_4$ provides the exact continuous representation for the Henderson filters, and, when $m = 6$, it results

$$K_4(t) = \frac{45}{7199963584} \left(-51619 + 156541\, t^2\right) \left(-1 + t^2\right) \left(-64 + 49\, t^2\right) \left(-81 + 49\, t^2\right).$$

The main drawback of this formulation is that it depends on the length of the filter, and it needs to be calculated any time that $m$ changes.

The asymptotic equivalent kernel representation based on the triweight density function allows to overcome such limitation, but Dagum and Bianconcini (2008) derived an equivalent reproducing kernel representation based on the biweight density function $f_{0B}(t) = \frac{15}{16}(1-t^2)^2, t \in [-1,1]$. In particular, Bianconcini and Quenneville (2010) showed that when the filter length is rather short, such as 9, 13 and 23 terms, the biweight density function is more appropriate to approximate the weight diagram of the Henderson filter. On the other hand, for longer filters, the triweight density function is a better choice.

## 3. Optimal real time filters

The derivation of the two-sided symmetric Henderson filter has assumed the availability of $2m+1$ observations centered at $t$. Obviously, for a given finite sequence $\{y_t, t = 1, ..., T\}$, it is not possible to obtain the estimates of the signal for the (first and) last $m$ time points. This is a serious limitation since, in real time analysis, we are interested in the most recent estimates. Several approaches have been followed in the literature to derive the asymmetric filters associated to the symmetric Henderson average.
As a local polynomial estimator, they have been obtained by fitting a local polynomial to the available observations $y_t, t = T-m+1, ..., T$, that is by minimizing

$$\sum_{j=-m}^{q} W_j (y_{t+j} - \beta_0 - \beta_1 j - ... - \beta_p j^p)^2, \quad q = 0, ..., m-1,$$

where $W_j$ denote the weighting function given in eq. (5). The corresponding weights satisfy the cubic polynomial reproduction constraints as in the interior of the time support, but the preservation of this unbiasedness property is done at the expenses of increasing the variance, which is very high, since most of the contribution to the trend estimate comes from the current observation.

In the X11ARIMA seasonal adjustment procedure, Dagum (1982) proposed to apply the symmetric two-sided filter to the series extended by $m$ forecasts. This is a safer approach provided that we are capable of producing optimal forecasts by means of some parametric or nonparametric model (*e.g.* by fitting a time series model of the ARIMA class). An intuitive and easily established fact is that if the forecasts are optimal in the mean square error sense, then the variance of the revision is minimum (see Wallis, 1983).

A different and wider applied method can be found in Musgrave (1964). This technique is used in the X-11 to derive the end weights of the trend-cycle averages corresponding to the Henderson filter. This approach assumes that the last data points of the time series follow a simple linear trend. Then the asymmetric weights $\{v_{-m}, ..., v_q\}$ are derived to minimize the expected square revision error

$$E\left(\sum_{j=-m}^{q} v_j y_j - \sum_{j=-m}^{m} w_j y_j\right)^2$$

under the constraint that $\sum_{j=-m}^{q} v_j = 1$. A complete and detailed review on this method can be found in Doherty (2001) and Quenneville et al. (2003). In particular, the latter showed that these filters are equivalent to apply the Henderson symmetric average on the available $m+q+1$ data points with the missing last $m-q$ points replaced by forecasts obtained as linear combinations of the first available $m+q+1$ data points.

### 3.1 Boundary kernels

A different derivation of the asymmetric Henderson filters can be provided by taken into account the reproducing kernel representation of the filter as introduced in the previous

Section. The Henderson filter is transformed into a continuous kernel function, and, at each time point $t^*$, the estimate $\hat{y}_{t^*}$ is obtained as a weighted mean of the sample $\{y_t, t = 1, ..., T\}$ with observations close to $t^*$ receiving the largest weights. The latter are provided by the kernel function $K$ defined on $[-1, 1]$, and specified only up to an unknown smoothing parameter $b$. That is (Nadaraya, 1964; Watson, 1964),

$$\hat{y}_t = \frac{\sum_{t=1}^{T} K\left(\frac{t-t^*}{b}\right) y_t}{\sum_{t=1}^{T} K\left(\frac{t-t^*}{b}\right)} \cong \frac{\sum_{j=-m}^{m} K\left(\frac{j}{b}\right) y_{t+j}}{\sum_{j=-m}^{m} K\left(\frac{j}{b}\right)}. \qquad (10)$$

The equivalence between the two formulations in eq. (10) derives from the fact that the bandwidth $b$ is selected to ensure that only $2m + 1$ observations surrounding the target point will receive nonzero weights.
The approach is quite simple, but presents some flaws. Indeed, at the boundary of the predictor space ($t = 1, ..., m$, and $t = T-m+1, ..., T$), the kernel neighborhood is asymmetric ($j = -m, ..., q; q = 0, ..., m-1$) and the estimates may be biased and show an increased variance. Therefore, a variety of kernel modifications have been proposed in the literature to provide approximated and asymptotic adjustments to overcome these drawbacks.
In order to derive the asymmetric weights associated to the symmetric Henderson average, Dagum and Bianconcini (2008, and 2011) applied the "cut and normalize" method proposed by Gasser and Muller (1979). The latter showed that, for a fixed symmetric filter of length $2m + 1$, and corresponding bandwidth $b$, the effective domain of the kernel function in the boundaries is $[-1, q^*)$, where $q^* = q/b$, instead of $[-1, 1]$ as for any interior point. For simplicity, only the left boundary effects, *i.e.* $q^* < 1$, will be discussed here. The right boundary effects proceed in the same manner. Since $s = j/b \in [-1, q^*)$, the symmetry of the kernel is lost, such that $\int_{-1}^{q^*} K(s)ds \neq 1$ and $\int_{-1}^{q^*} s^i K(s)ds \neq 0, i = 1, .., p-1$.
The boundary kernel $K_{q^*}$ is obtained by "cutting" the symmetric kernel $K$ to omit that part of the function lying between $q^*$ and 1, and by "normalizing" it between -1 and $q^*$ as follows

$$K_{q^*}(s) = \frac{K(s)}{\int_{-1}^{q^*} K(s)dt}. \qquad (11)$$

Boundary kernels (11) satisfy the following properties (Gasser and Muller, 1979)

**1. Moment conditions:** $\int_{-1}^{q^*} K_{q^*}(s)ds = 1, \int_{-1}^{q^*} K_{q^*}(s)s^i ds \neq 0, i = 1, ..., p.$

**2. Variance condition:** $\int_{-1}^{q^*} K_{q^*}^2(s)ds < c, \forall q^* \in [0, 1)$, that is the asymptotic variance has to be uniformly bounded.

**3. Convergence:** the kernels depends continuously on $q^*$, and

$$K_{q^*} \xrightarrow[q^* \to 1]{} K.$$

Property 3. is particularly appealing when asymmetric kernels are applied for trend-cycle estimation. In fact, a monotonic convergence to the corresponding symmetric filter should imply a reduction of the revisions when new observations are added to the series. On the other hand, Property 1. shows that these kernels are biased estimators of polynomial trend of degree $p$ at the boundaries. However, since the neighborhood of observations is smaller than in the interior of the support, this can be allowed in order to reduce the variability of the estimates as shown by Property 2.

Boundary Henderson kernels can be derived starting from the exact, the biweight and the triweight third order kernel representations, as provided before. However, the exact density function depends on the length of the filter, hence we now concentrate on the study

of the boundary behavior of the biweight and triweight kernels. At this regard, Bianconcini and Quenneville (2010) showed that their convergence to the corresponding symmetric kernel ($q^* = 1$) is quite fast (in general after the previous to the last point asymmetric filter).
The application of the "cut and normalize" method brings the following formula for the asymmetric weights

$$w_j = \frac{K(j/b)}{\sum_{j=-m}^{q} K(j/b)} \qquad j = -m, ..., q; q = 0, ..., m-1 \tag{12}$$

where $K$ is the symmetric kernel function, and $b$ is the bandwidth parameter that can be different for each of the $m$ asymmetric averages.
The choice of the kernel function $K$ is based on the properties of the corresponding boundary kernels. First of all, this requires to select the hierarchy of kernels and, then, within each hierarchy the most appropriate order for the kernel function. Coherently with the symmetric weights, the biweight and the triweight represent good alternative hierarchies for deriving asymmetric Henderson filters for short lengths and asymptotically, respectively. Furthermore, to preserve the monotonic convergence to the symmetric filters, fourth order kernels are also applied at the boundaries of the sample. The selection of the bandwidth parameter $b$ is a crucial choice. In practice, for asymmetric filters, these parameters are chosen a priori in order to ensure a prespecified filter length ($m + q + 1, q = 0, ..., m - 1$), as well as a rapid convergence to the corresponding symmetric filter. In the following, we discuss the choice of optimal kernels and bandwidth parameters, optimal in the sense of minimizing the size of revisions in the final estimates.

### 3.2 A measure of filter revisions

Following Dagum (1982), to derive a measure of filter revision, the behavior of the different smoothers is analyzed with respect to cyclical components. The set of weights $\{w_j, j = -m, ..., q\}$ is called the *impulse response function* of the filter, and its properties are described in the frequency domain by looking at its Fourier transform, $H(\omega)$, called the *frequency response function*. That is,

$$\begin{aligned} \hat{y}_t &= \sum_{j=-m}^{q} w_j y_{t+j} = \sum_{j=-m}^{q} w_j \exp(-i\omega(t+j)) \\ &= H(\omega)\exp(-i\omega t) = H(\omega) y_t. \end{aligned}$$

$H(\omega)$ fully describes the effect of the linear filter on the given input. In particular,

$$H(\omega) = G(\omega)\exp(-i\phi(\omega)) \tag{13}$$

where $G(\omega)$ is called the *gain function* of the filter and $\phi(\omega)$ is the *phase shift function*. $G(\omega)$ describes how much the amplitudes of a time series are amplified or reduced at the frequency $\omega$, whereas $\phi(\omega)$ contains information about the translation on time of the cyclical components.

In order to analyze and compare the properties of the boundary asymmetric kernels, we take into account the frequency response function, that incorporates both the information coming from the gain function in terms of revisions, and the information coming from the phaseshift function in terms of delay to detect true turning points. Hence, we compute the Euclidean distance between the frequency response function of the asymmetric filter $H_a(\cdot)$ and the frequency response function of the corresponding symmetric Henderson filter $H(\cdot)$, that is

$$\sqrt{\int_0^{1/2} |H_a(\omega^*) - H(\omega^*)|^2 d\omega^*} \quad (14)$$

In order to improve the convergence of the boundary filters, optimal bandwidths for each asymmetric filter, can be selected in order to minimize eq. (14).
Table 1 illustrates the optimal values selected for each of the biweight and triweight boundary filters corresponding to the 9- and 13-term biweight and triweight symmetric weights, the latter computed by fixing the bandwidths to $m + 1$ (see for details Bianconcini and Quenneville, 2010). We can notice that the bandwidths are quite close for the several asymmetric filters, with values ranging from $m$ to $m + 1$.

| $q$ | 0 | 1 | 2 | 3 | | |
|---|---|---|---|---|---|---|
| Biweight | 4.01 | 4.23 | 5.13 | 4.13 | | |
| Triweight | 4.01 | 4.69 | 4.01 | 4.28 | | |
| $q$ | 0 | 1 | 2 | 3 | 4 | 5 |
| Biweight | 6.01 | 6.01 | 6.38 | 6.01 | 6.01 | 6.50 |
| Triweight | 6.01 | 6.01 | 6.78 | 6.01 | 6.01 | 6.67 |

**Table 1**: Optimal bandwidths for the biweight and triweight boundary kernels associated to 9-term and 13-term biweight and triweight optimal symmetric filter.

Looking at the spectral properties of the related asymmetric weights, Table 2 shows that the convergence of the filters to the associated symmetric ones is improved mainly for the last point filter, and for $q \geq 1$ the distances are almost null.

| $q$ | 0 | 1 | 2 | 3 | | |
|---|---|---|---|---|---|---|
| Biweight | 0.06 | 0.01 | 0.02 | 0.00 | | |
| Triweight | 0.07 | 0.01 | 0.03 | 0.00 | | |
| $q$ | 0 | 1 | 2 | 3 | 4 | 5 |
| Biweight | 0.12 | 0.03 | 0.02 | 0.03 | 0.01 | 0.00 |
| Triweight | 0.14 | 0.05 | 0.01 | 0.03 | 0.02 | 0.00 |

**Table 2**: Euclidean distances between the frequency response function of the biweight and triweight boundary kernels and the frequency response function of the corresponding 9-term and 13-term symmetric filters.

The asymmetric weights based on optimal bandwidths do not satisfy the unbiasedness property of the associated symmetric averages. However, these filters present values of the first moment similar to the Musgrave filters. Hence, when optimal bandwidths are selected, the corresponding weights could pass a linear trend at the end of the sample period with small bias, mainly for the first and second asymmetric filter. Similar conclusions can be drawn for asymmetric weights associated to symmetric filter of any length, not reported here for space reasons.
On the other hand, looking at the variability of the estimates, the leverage values of these filters are computed. For the 9-term symmetric filter, the weight associated to the current observation for the last point asymmetric filter is equal to 0.58 and 0.63 for the biweight and triweight kernels, respectively. On the other hand, for the 13-term filter, these weights are equal to 0.43 and 0.47 for the biweight and triweight, respectively.
In order to better evaluate the relationship of these optimal asymmetric weights with the Musgrave filters, we compute their Euclidean distances as shown in Table 4. We notice that the spectral properties of these filters are very similar, and we can claim that they provide a very good compromise in terms of both revisions (gain function) and detection of true turning points (phaseshift function).

| $q$ | 0 | 1 | 2 | 3 | $q$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|---|---|---|---|---|
| $\sum_{j=-m}^{m} j w_j$ | | | | | $\sum_{j=-m}^{m} j^2 w_j$ | | | | |
| Biweight | -0.35 | 0.02 | 0.16 | 0.01 | Biweight | 0.06 | 0.33 | 0.81 | 0.16 |
| Triweight | -0.28 | 0.03 | 0.12 | 0.01 | Triweight | 0.01 | 0.28 | 0.34 | 0.06 |
| $\sum_{j=-m}^{m} j^3 w_j$ | | | | | $\sum_{j=-m}^{m} j^4 w_j$ | | | | |
| Biweight | 0.98 | 1.18 | 2.56 | 0.17 | Biweight | -4.44 | -4.06 | -5.82 | -6.22 |
| Triweight | 0.94 | 1.40 | 1.07 | 0.12 | Triweight | -3.81 | -5.07 | -2.13 | -5.95 |

**Table 3**: Properties of optimal asymmetric filters associated with the symmetric 9-term Henderson filter.

| | $m$ | |
|---|---|---|
| *Kernels* | 4 | 6 |
| Biweight | 0.19 | 0.13 |
| Triweight | 0.23 | 0.17 |

**Table 4**: Euclidean distances of the optimal last point biweight and triweight weights and last point Musgrave filter for different values of $m$.

## 4. Empirical application

This section examines how the fourth order biweight and triweight boundary kernels perform on real data in comparison with the classical asymmetric Henderson filters developed by Musgrave (1964).

We apply the last point filters to a set of 30 time series taken from the Hydman's time series library (http://www-personal.buseco.monash.edu.au/ hyndman/TSDL/). These series are seasonally adjusted and related to different fields (finance, labor market, production, sales, transports and tourism, utilities). The periods selected vary to sufficiently cover the various lengths published for these series. We want to study how the kernels and classical estimators respond to the variability of the data. For each series, the length of the symmetric filters is selected according to the $I/C$ (noise to signal) ratio, as classically done in the X11/X12ARIMA procedure. In the sample, the ratio ranges from 0.16 to 7.35, hence filters of length 9, 13, and 23 terms are applied. The comparisons are based on the relative filter revisions between the final symmetric filter $S$ and the last point asymmetric filter $A$, that is,

$$R_t = \frac{S_t - A_t}{S_t}, \quad t = 1, 2, ..., N. \tag{15}$$

For each series and for each estimator, we calculate the ratio between the Mean Square Error (MSE) of the revisions corresponding to the fourth order kernel $R^K$ and to the corresponding classical filter $R^C$, that is $\frac{MSE(R^K)}{MSE(R^C)}$. A comparison is also performed between the biweight and triweight boundary kernels in order to analyze how they differently behave on real data. For all the estimators, the results, illustrated in Table 5, indicate that the kernel last point predictors introduce smaller revisions than the classical ones, being the ratio always smaller than one. This implies that the kernel estimates will be more reliable and efficient than the ones obtained by the application of the classical Henderson filter. In particular, for the biweight estimator, 80% of the sample has a ratio less than 0.93 and, in general, it is never greater than 0.96. For the triweight filter, the ratio is less than 0.93 in 57% of the series, with highest value equal to 0.99. Furthermore, as highlighted by the last column of Table 5, the biweight filter generally performs better than the triweight kernel, even if their performance is very similar for all the series in the sample.

For current economic analysis, it is also of great importance to assess the performance of the several estimators for turning point detection. A turning point is generally defined to

**Table 5**: MSE revision ratios between kernels and classical last point predictors, as well as between boundary kernels.

| Macro-area | Series | Biweight/Classical | Triweight/Classical | Triweight/Biweight |
|---|---|---|---|---|
| Finance | Monthly return on the S&P 500 index | 0.85 | 0.89 | 1.05 |
| | Return to an investment strategy based on the paper rate | 0.93 | 0.96 | 1.03 |
| | Commercial paper rate | 0.85 | 0.88 | 1.03 |
| | Railroad bond yields | 0.84 | 0.84 | 1.00 |
| | Mutual savings bank data end-of-month balance | 0.82 | 0.73 | 0.89 |
| Labour Market | Wisconsin employment, fabricated metals | 0.76 | 0.75 | 0.98 |
| | US male (20 years and over) unemployment figures | 0.93 | 0.96 | 1.04 |
| | Unemployment benefits Australia | 0.93 | 0.96 | 1.02 |
| | Women unemployed UK | 0.86 | 0.89 | 1.04 |
| | Canadian total unemployment figures | 0.84 | 0.83 | 0.99 |
| | US female (20 years and over) unemployment rate | 0.84 | 0.86 | 1.01 |
| | Number of employed persons in Australia | 0.91 | 0.95 | 1.04 |
| Production | Basic iron production in Australia | 0.91 | 0.95 | 1.05 |
| | Production of chocolate confectionery in Australia | 0.94 | 0.98 | 1.04 |
| | Production of Portland cement | 0.93 | 0.93 | 1.00 |
| | Electricity production in Australia | 0.93 | 0.97 | 1.05 |
| | Production of blooms and slabs in Australia | 0.95 | 0.89 | 0.94 |
| | Production of blooms and slabs | 0.94 | 0.93 | 0.99 |
| Sales | Sales of tasty Cola | 0.83 | 0.81 | 0.97 |
| | Unit sales, Winnebago Industries | 0.93 | 0.95 | 1.03 |
| | Sales of new one-family houses sold in US | 0.93 | 0.97 | 1.04 |
| | Sales of souvenir shops in Quennsland, Australia | 0.96 | 0.99 | 1.03 |
| | Demand for carpet | 0.96 | 0.99 | 1.03 |
| Transport and Tourism | Portland Oregon, average bus ridership | 0.83 | 0.80 | 0.96 |
| | US air passenger miles | 0.91 | 0.91 | 0.99 |
| | International airline passengers | 0.84 | 0.84 | 1.00 |
| | Weekday bus ridership, Iowa city | 0.93 | 0.93 | 1.00 |
| | Passenger miles flow domestic UK | 0.93 | 0.93 | 1.00 |
| Utilities | Average residential gas usage Iowa | 0.93 | 0.98 | 1.06 |
| | Total number of customers | 0.94 | 0.99 | 1.05 |

occur at time $t$ if (*downturn*):

$$y_{t-k} \leq \cdots \leq y_{t-1} > y_t \geq y_{t+1} \geq \cdots \geq y_{t+m}$$

or (*upturn*)

$$y_{t-k} \geq \cdots \geq y_{t-1} < y_t \leq y_{t+1} \leq \cdots \leq y_{t+m}$$

for $k = 3$ and $m = 1$ given the smoothness of the trend cycle data (see, for more details, Dagum, 1996). For each estimator, the time lag to detect the true turning point is obtained by calculating the number of months it takes for the revised trend series to signal a turning point in the same position as in the final trend series.

True turning points are those recognized in the economy after occurring. For instance, the Business Cycle Dating Committee of the National Bureau of Economic Research (NBER) had identified the last trough in the US economy in June 2009. To determine this date, the behavior of various monthly indicators of the economic activity has been analyzed. Among others, the committee considered a measure of monthly Gross Domestic Product (GDP) developed by the private forecasting firm Macroeconomic Advisers (MAGDP), monthly GDP and Gross Domestic Income (GDI) developed by James H. Stock and Mark W. Watson (SWGDP, SWGDI), real personal income excluding transfers (INC), the payroll (PAY) and household measures of total employment (HOU). Monthly data series for industrial production (IIP) and manufacturing trade sales (MAN) have been also taken into account. The committee designated June as the month of the trough for several of these indicators, that is MAGDP, SWGDP, SWGDI, MAN, IIP. On the other hand, the trough was identified in October 2009 for INC, and in December 2009 for PAY and HOU. The results obtained by applying the classical and kernel estimators to all of these series are shown in Table 6.

It can be noticed that kernel asymmetric filters are quicker to detect turning points than the Musgrave filters. In particular, the biweight kernel always detects the turning points

**Table 6**: Time lag in detecting true Turning Points (TP) for Musgrave and kernel estimators.

| Series | True TP | Length | Musgrave | Biweight | Triweight |
|---|---|---|---|---|---|
| MAGDP | 06/09 | 13 | 5 | 5 | 4 |
| SWGDP | 06/09 | 13 | 5 | 5 | 4 |
| SWGDI | 06/09 | 13 | 5 | 1 | 4 |
| MAN | 06/09 | 9 | 1 | 2 | 1 |
| IIP | 06/09 | 9 | 2 | 2 | 1 |
| INC | 10/09 | 13 | 5 | 1 | 4 |
| PAY | 12/09 | 9 | 4 | 2 | 1 |
| HOU | 12/09 | 9 | 1 | 2 | 1 |

only very few months after it has occurred, and faster than the classical Musgrave filter, which is very well known for fast turning point detection. It also performs similarly to the triweight kernel, but in some cases it detects the true turning point with much shorter delay, that is three months earlier than the triweight. As highlighted in Table 7, for these series, the biweight kernel also behaves better than the classical and the triweight in terms of total revisions.

**Table 7**: Total size revisions for the NBER series.

| Series | Bi/Cl | Tri/Cl | Tri/Bi |
|---|---|---|---|
| MAGDP | 0.926 | 0.936 | 1.011 |
| SWGDP | 0.926 | 0.936 | 1.011 |
| SWGDI | 0.921 | 0.931 | 1.010 |
| MAN | 0.851 | 0.896 | 1.053 |
| IIP | 0.857 | 0.905 | 1.056 |
| INC | 0.926 | 1.016 | 1.097 |
| PAY | 0.787 | 0.996 | 1.266 |
| HOU | 0.777 | 0.994 | 1.280 |

## 5. Conclusions

In this study, we have derived asymmetric trend-cycle filters, which are applied to the most recent observations in order to obtain real time estimates.

We made use of the RKHS methodology, according to which hierarchies of Henderson kernels are generated via the multiplication of the biweight and triweight density functions with the corresponding orthonormal polynomials. These kernels depend on bandwidth parameters, that we selected using an optimal criteria to reduce revisions in the final estimates, and to improve the detection of true turning points.

We found that the optimal biweight boundary kernels perform generally better than the triweight ones both in terms of revisions and detection of true turning points. Furthermore, both the boundary kernels always perform better than the classical Henderson filters applied for real time estimation as developed by Musgrave (1964). In particular, the latter are very well-know for fast turning point detection, but we have shown that the boundary kernels can detect turning points with much shorter delay than the Musgrave filters, sometimes more than three months in advance.

## REFERENCES


Berlinet, A. and Thomas-Agnan, C. (2003), *Reproducing Kernel Hilbert Spaces in Probability and Statistics*, Kluwer Academic Publishers.

Brezinski, C. (1980), *Pade Approximation and General Orthogonal Polynomials*, Birkhauser, Basel.

Bianconcini S. and Quenneville B. (2010), Real time analysis based on reproducing kernel Henderson filters, *Estudios de Economia Aplicada*, 28(3), 553-574.

Dagum, E.B. (1982), The effects of asymmetric filters of seasonal factor revisions, *Journal of the American Statistical Association*, 77, 732-738.

Dagum, E.B. (1988), The X11ARIMA/88 Seasonal Adjustment Method - Foundation and User's Manual, Research Paper, Time Series Research and Analysis Division, Statistics Canada, Ottawa.

Dagum, E.B. (1996), "A New Method to Reduce Unwanted Ripples and Revisions in Trend-Cycle Estimates from X11ARIMA," *Survey Methodology*, 22, 77-83.

Dagum, E.B. and Bianconcini, S. (2008), The Henderson Smoother in Reproducing Kernel Hilbert Space, *Journal of Business and Economic Statistics*, 26 (4), 536-545.

Dagum, E.B. and Bianconcini, S. (2011), A unified probabilistic view of nonparametric predictors via reproducing kernel Hilbert spaces, *Econometric Reviews*, to appear.

Doherty, M. (2001), Surrogate Henderson filers in X-11, *Australian and New Zealand Journal of Statistics*, 43,385-392.

Fan, J. and Gijbels, I. (1996), *Local polynomial modelling and its applications*, Monographs on statistics and applied probability 66, Chapman and Hall/CRC.

Gasser, T. and Muller, H.G. (1979). Kernel Estimation of Regression Functions. In *Smoothing Techniques for Curve Estimation*, Lecture Notes in Mathematics 757, eds. T. Gasser and M. Rosenblatt, Heidelberg: Springer-Verlag, pp. 23-68.

Gray, A. and Thomson, P. (1996), Design of Moving-Average Trend Filters Using Fidelity and Smoothness Criteria, in *Time Series Analysis* (in memory of E.J. Hannan), eds. Robinson, P.M. and Rosenblatt, M., vol.II, New York: Springer Lecture Notes in Statistics, 115, 205-219.

Henderson, R. (1916), Note on Graduation by Adjusted Average, *Transaction of Actuarial Society of America*, 17, 43-48.

Kenny, P. and Durbin (1982), Local Trend Estimation and Seasonal Adjustment of Economic and Social Time Series, *Journal of the Royal Statistical Society A*, 145, 1-41.

Loader, C. (1999), *Local Regression and Likelihood*, New York: Springer.

Muller, H.G. (1984), Smooth Optimum Kernel Estimators of Regression Curves, Densities and Modes, *Annals of Statistics*, 12, 766-774.

Musgrave, J. (1964), A set of end weights to end all end weights, Working paper. Washington DC: Bureau of Census.

Nadaraya, E. A. (1964). "On Estimating Regression". *Theory of Probability and its Applications*, 9 (1): 141-142.

Quenneville, B. and Ladiray, D. (2000), Locally adaptive trend-cycle estimation for X11, *Proocedings of the International Conference on Establishment Surveys-II*, Buffalo, NY, June 2000.

Quenneville, B., Ladiray, D. and Lefrancois, B. (2003), A note on Musgrave asymmetrical trend-cycle filters, *International Journal of Forecasting*, 19, 727-734.

Wallis, K. (1983), Models for X-11 and X-11 forecast procedures for preliminary and revised seasonal adjustments, in *Applied Time Series Analysis of Economic Data*, A. Zellner ed., Washington DC: Bureau of the Census, 3-11. 1346-1370.

Watson, G. (1964). Smooth regression analysis. *Sankhya Series, A*, 26, 359-372.