LOESS Asymmetric Filters for Real Time Economic Analysis

Silvia Bianconcini

Department of Statistics, University of Bologna

e-mail: bianconcini@stat.unibo.it.

Abstract

The local polynomial regression predictors developed by Cleveland (1979) have been applied to estimate the short-term trend of seasonally adjusted economic indicators and implemented in several statistical packages. The main purpose of this study is to introduce a RKHS representation of the LOESS smoother with particular emphasis on the asymmetric filters applied to most recent observations. The filters obtained by means of the RKHS are shown to have superior properties relative to the classical ones from the view point of signal passing, noise suppression and revisions. We compare the performance of the kernel representations relative to the classical filters using real life series.

Keywords: Local polynomial regression, reproducing kernel, spectral properties, revisions, short term trend estimators.

JEL Classification: C14, C22.

1 Introduction

The basic approach to the analysis of current economic conditions, known as recession and recovery analysis (see Moore, 1961) is that of assessing the short-term trend of major economic indicators (leading, coincident and lagging) using percentage changes, based on original units calculated for months and quarters in chronological sequence. The main goal is to evaluate the behavior of the economic indicators during incomplete phases by comparing current contractions or expansions with corresponding phases in the past. This is done by measuring changes of single time series (mostly seasonally adjusted) from their standing at cyclical turning points with past changes over a series of increasing spans. In recent years, statistical agencies have shown an interest in providing trend-cycle estimates or smoothed seasonally adjusted data to facilitate recession and recovery analysis. Among other reasons, this interest originated from major economic and financial changes of global nature which have introduced more variability in the data, and consequently in the seasonally adjusted numbers, making very difficult to determine the direction of the short-term trend, particularly to assess the presence or the upcoming of a turning point.

The estimation of the trend-cycle with the X11/X12ARIMA seasonal adjustment method (see Dagum, 1988 and Findley et al., 1998) as well as the US Bureau of Census method II - X11 variant (Shiskin et al., 1967) is done by the application of linear filters due to Henderson (1916). These estimators possess similar (fitting and smoothing) properties to the local polynomial regression predictors developed by Cleveland (1979), as noted and studied by Dagum and Bianconcini (2006a and 2006b).

The LOESS filters have been applied to estimate the short-term trend of seasonally adjusted economic indicators in the STL procedure (Cleveland et al., 1990), and implemented in the most widely used statistical packages, such as Eviews, Stata, S-plus, R, Matlab, SAS.

The LOESS estimator, originally called LOWESS (LOcally WEighted Scatterplot Smoother), is based on nearest neighbor weights and is applied in an iterative manner for robustification. LOESS of order p consists of locally fitting a p-th degree polynomial by means of weighted least squares. The weighting function suggested and chosen by Cleveland (1979) is the tricube one. The asymmetric weights of the filters are derived following the same technique by weighting the data belonging to an asymmetric neighborhood which contains the same number of data points of the symmetric one.

The study of the properties and limitations of this smoother have been discussed by many authors, among them, Kenny and Durbin (1982), Cleveland and Devlin (1988), Cleveland et al.(1990), Wand and Jones (1995), Gray and Thomson (1990 and 1996), Fan and Gijbels (1996), Dagum and Luati

(2000 and 2004). However, to the best of our knowledge, none of these studies have approached the LOESS filter from a Reproducing Kernel Hilbert Space (RKHS) perspective.

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.
The RKHS approach followed in our study is strictly nonparametric. We make use of the fundamental theoretical result due to Berlinet (1993) according to which a kernel estimator of order $p \geq 2$ can always be decomposed into the product of a reproducing kernel $R_{p-1}(\cdot, \cdot)$ and a probability density function f_0 with finite moments up to order $2p$. The reproducing kernel belongs to the space of polynomials of degree at most $p - 1$.

The main purpose of this study is to introduce a RKHS representation of the LOESS smoother with particular emphasis on the asymmetric ones applied to most recent observations. The asymmetric filters can be derived coherently with the corresponding symmetric weights or from a lower or higher order kernel within a hierarchy, if preferred. In the particular case of the currently applied asymmetric LOESS filters, those obtained by means of the RKHS are shown to have superior properties relative to the classical ones from the view point of signal passing, noise suppression and revisions. We compare the performance of the kernel representations relative to the classical filters using real life series.

Section 2 briefly presents the basic properties of reproducing kernel Hilbert spaces. Section 3 discusses the classical LOESS symmetric smoother and its kernel representation. Section 4 presents the asymmetric filters of LOESS and performs a spectral analysis comparison with those currently in use. Section 5 illustrates the new asymmetric LOESS kernel smoothers with applications to real data. Finally, Section 6 gives the conclusions.

2 Linear Filters in Reproducing Kernel Hilbert Spaces

Let $\{(t_i, y_i), i = 1, 2, ..., N\}$ denote each time point t_i at which the observation y_i is taken. Without loss of generality, we will assume that $0 \leq t_1 \leq t_2 \leq ... \leq t_N \leq 1$.

A basic assumption in time series analysis is that the input series $\{y_i, i = 1, 2, ..., N\}$ can be decomposed into the sum of a systematic component called the signal (or nonstationary mean) g_i, plus an erratic component called the noise u_i, such that

$$y_i = g_i + u_i. \tag{1}$$

The noise component u_i 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.

Assuming that the input series $\{y_i, i = 1, 2, ..., N\}$ is seasonally adjusted or without seasonality, the signal g_i 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 have a global or a local representation. Following a nonparametric regression approach, we only assume that the curve g_i belongs to some infinite dimensional function space, chosen to satisfy some smoothness (*i.e.* continuity and differentiability) conditions. Hence, g_i can be represented *locally* by a polynomial of degree p (using a p-th order Taylor approximation) of the time distance j, between y_i and the neighboring observations $y_{i-j}, j = -m, ..., m$. Hence, given u_i for some time point t_i, it is possible to find a local polynomial trend estimator

$$g_i(j) = a_0 + a_1 j + ... + a_p j^p + \varepsilon_i(j), \tag{2}$$

where $a_0, a_1, ..., a_p \in \mathbb{R}$ and ε_i is assumed to be purely random and mutually uncorrelated with u_i.

The coefficients $a_0, a_1, ..., a_p$ can be estimated by ordinary or weighted least squares or by summation formulae (see *e.g.* Dagum, 1985). The solution for $\hat{a}_0$ provides the trend-cycle estimate $\hat{g}_i(0)$, which equivalently is a weighted average applied in a moving manner (Kendall, Stuart, and Ord, 1983), such that

$$\hat{g}_i(0) = \hat{g}_i = \sum_{j=-m}^{m} w_j y_{i-j} \tag{3}$$

where $w_j, j < N$, denotes the weights to be applied to the observations y_{i-j} to get the estimate $\hat{g}_i$ for each point in time $t_i, i = 1, 2, ..., N$.

The weights depend on: (1) the degree of the fitted polynomial, (2) the amplitude of the neighborhood, and (3) the shape of the function used to average the observations in each neighborhood.

Once a (symmetric) span $2m + 1$ of the neighborhood has been selected, the w_j's for the observations corresponding to points falling out of the neighborhood of any target point are null or approximately null, such that the estimates of the $N - 2m$ central observations are obtained by applying $2m + 1$ symmetric weights to the observations neighboring the target point. The missing estimates for the first and last m observations can be obtained by applying asymmetric moving averages of variable length to the first and last m observations, respectively. The length of the moving average or time invariant symmetric linear filter is $2m+1$, whereas the asymmetric linear filters length is time varying.

Using the backshift operator B, such that $B^n y_t = y_{t-n}$, equation (3) can be written as

$$\hat{g}_i = \sum_j w_j B^j y_i = W(B) y_i, \quad i = 1, 2, ..., N \tag{4}$$

where $W(B)$ is a linear nonparametric estimator.

The nonparametric estimator $W(B)$ is said to be a second order kernel if it satisfies the conditions

$$\sum_{j=-m}^{m} w_j = 1, \tag{5}$$

$$\sum_{j=-m}^{m} j w_j = 0, \tag{6}$$

hence it preserves a constant and a linear trend. On the other hand, $W(B)$ is a higher order kernel if

$$\sum_{j=-m}^{m} w_j = 1, \tag{7}$$

$$\sum_{j=-m}^{m} j^r w_j = 0, \tag{8}$$

for some $r = 1, 2, ..., p \geq 2$. In other words, it will reproduce a polynomial trend of degree $p-1$ without distortion.

A different characterization of a p-th order nonparametric estimator can be provided by means of the RKHS methodology.

A Hilbert space $\mathcal{H}$ is a complete linear space with a norm given by an inner product. The space of square integrable functions L^2 and the finite p-dimensional space $\mathbb{R}^p$ are those used in this study.

Let us assume that the time series $\{y_i, i = 1, 2, ..., N\}$ is a finite realization of a family of square integrable random variables, *i.e.* $\int_0^1 |Y_t|^2 dt < \infty$. Hence, the stochastic process $\{Y_t\}_{t\in[0,1]}$ belongs to the space $L^2[0,1]$.

The space $L^2[0,1]$ is a Hilbert space endowed with the inner product defined by

$$< U(t), V(t) >= \int_0^1 U(t)V(t)f_0(t)dt$$

where $U(t), V(t) \in L^2[0,1]$ and $f_0(t)$ is a continuous design density on the bounded interval $[0,1]$. In the following, $L^2[0,1]$ will be indicated as $L^2(f_0)$. The local trend $g_i(\cdot)$ belongs to the space of polynomials of degree at most p, $\mathbf{P}_p$, with p non-negative integer. $\mathbf{P}_p$ is a Hilbert subspace of $L^2(f_0)$, hence it inherits its inner product, such that

$$< P(t), Q(t) >= \int_0^1 P(t)Q(t)f_0(t)dt, \tag{9}$$

where $P(t), Q(t) \in \mathbf{P}_p$.

Berlinet (1993) showed that the space $\mathbf{P}_p$ is a reproducing kernel Hilbert space of polynomials on some domain T, that is, $\forall t \in [0,1]$ and $\forall P \in \mathbf{P}_p$, there exists an element $R_p(t,.) \in \mathbf{P}_p$, such that

$$P(t) =< P(\cdot), R_p(t,\cdot) > . \tag{10}$$

$R_p(t,\cdot)$ is called reproducing kernel. Formally, R is a function $R : [0,1] \times [0,1] \rightarrow \mathbb{R}$ that satisfies the following properties:

(1) $R(t,\cdot) \in \mathcal{H}, \forall t \in [0,1]$; and

(2) $< g(\cdot), R(t,\cdot) >= g(t), \quad \forall t \in [0,1]$ and $\forall g \in \mathcal{H}$.

This last condition is called the "reproducing property": the value of the function g at the point t is reproduced by the inner product of g with $R(t,\cdot)$.

The estimate $\hat{g}_i$ can be equivalently seen either as the projection of y_i on $\mathbf{P}_p$ or as a local weighted average of the observations for the discrete version of the filter given in eq. (4). The weights w_j are derived by a kernel function K of order $p+1$,

$$K_{p+1}(t) = R_p(t,0) f_0(t), \tag{11}$$

where p is the degree of the fitted polynomial. Eq. (11) results from a fundamental theorem proved by Berlinet (1993) which states that:

Kernels of order $(p+1)$, $p \geq 1$, can be written as products of the reproducing kernel $R_p(t,\cdot)$ of the space $\mathbf{P}_p \subset L^2[0,1]$ and a density function f_0 with finite moments up to order $2p$.

It follows that for any sequence $(P_l)_{0\leq l\leq p}$ of $(p+1)$ orthonormal polynomials in $L^2(f_0)$,

$$R_p(t,0) = \sum_{l=0}^{p} P_l(t)P_l(0) \tag{12}$$

and therefore,

$$K_{p+1}(t) = \sum_{l=0}^{p} P_l(t)P_l(0)f_0(t). \tag{13}$$

An important outcome of the RKHS theory is that linear filters can be grouped into hierarchies $\{K_p, p = 2,3,4,...\}$ with the following property: each hierarchy is identified by a density f_0 and contains kernels of order 2, 3, 4,... which are products of orthonormal polynomials by f_0.

The weight system of a hierarchy is completely determined by specifying: (a) the bandwidth or smoothing parameter, which determines the length of the filter, (b) the maximum order of the estimator in the family, and (c) the density f_0.

In this study, the smoothing parameter is not derived by data dependent optimization criteria, but we fixed it with the aim to obtain a kernel representation of the most often applied LOESS smoothers for current economic analysis. However, kernels of any length, including infinite ones, can be obtained with the above approach. Consequently, the results discussed can be easily extended to any filter length as long as the density function and its orthonormal polynomials are specified.

The identification and specification of the density is one of the most crucial task for smoothers based on local polynomial fitting by weighted least squares, as the LOESS filter. The density is related to the weighting function which compares in the least squares minimization problem.

3 LOESS Symmetric Smoothers and their Kernel Representation

Local polynomial fitting has a long history in time series, particularly within the actuarial literature. Henderson (1916), Whittaker and Robinson (1924), and Macauley (1931) are some of the earlier classical references. These authors were very much concerned with the smoothing properties of linear estimators, being Henderson the first to show that the smoothing power of a linear filter depends on the shape of its weighting system. Significant modern contributions include Cleveland (1979), Cleveland and Devlin (1988), Wand

and Jones (1995), Fan and Gijbels (1996), and Loader (1999).

The regression smoother, here considered and studied, is the one developed by Cleveland (1979), known as LOESS and originally called LOWESS (LOcally Weighted Scatterplot Smoother), based on the nearest neighbor weights and applied in an iterative manner for robustification.

Given a series of equally spaced observations and corresponding target points $\{(y_i, t_i), i = 1, ..., N\}$, $0 \leq t_1 \leq ... \leq t_N \leq 1$, LOESS produces a smoothed estimate $\hat{g}_i$ as follows

$$\hat{g}_i = \mathbf{t}_i^T \hat{\beta}_i \tag{14}$$

where $\mathbf{t}_i$ is a $(p+1)$-dimensional vector of generic component $t_i^d, d = 0, ..., p$, $p = 0, 1, 2, ...$ denotes the degree of the fitted polynomial, and $\hat{\beta}_i$ is the $(d+1)$-dimensional least squares estimate of a weighted regression computed over a neighborhood of t_i constituting a subset of the full span of the series.

The weights of the regression depend on the distance between the target point t_i and any other point belonging to its neighborhood, $N(t_i)$, through a weight function $W(t)$, such that

$$w_k(t_i) = W\left(\frac{|t_i - t_k|}{\Delta(t_i)} \right), \quad \forall t_k \in N(t_i), \tag{15}$$

where $\Delta(t_i) = \max_{t_k \in N(t_i)} |t_i - t_k|$. Each neighborhood is made of the same number of points chosen to be nearest to t_i, and the ratio between the amplitude of the neighborhood, say k, and the full span of the series, N, defines the bandwidth or smoothing parameter. It is sensible to choose an odd value for k in order to allow symmetric neighborhood for central observations.

Three are the crucial choices to be made in LOESS: (1) the shape of the weighting function $W(t)$, (2) the degree of the fitting polynomial p, and (3) the value of the smoothing parameter.

The weighting function we use is the tricube proposed by Cleveland et al. (1990), *i.e.*

$$W(t) = (1 - |t|^3)^3 I_{[-1,1]}(t), \tag{16}$$

whose quasi-semicircular shape allows about 45% of the points belonging to any symmetric neighborhood to have considerable weight (greater than 0.8), the remaining 55% having weights decreasing to zero quite slowly.

Concerning the degree of the fitting polynomial, $p = 1$ or $p = 2$ are usually appropriate choices. The highest degree is more appropriate when the plot of the flexibility of a quadratic curve best fits highly noisy time series.

For the smoothing parameter selection, Cleveland et al. (1988 and 1990) proposed some data dependent criteria, such as the diagnostic M-plot and the C_p procedure (Mallows, 1966 and 1973), whereas, in the iterative STL procedure, suggested to choose a filter length equal to the smallest odd integer greater than or equal to the periodicity of the data.

On the other hand, we fix the smoothing parameter to ensure filter lengths equal to 9, 13, and 23 terms, which are the most widely applied for the current economic analysis of monthly time series. We use the I/C ratio estimated on the series to specify the length of the trend-cycle moving average adopted. I and C are respective averages of the absolute values of the month to month changes in the (estimated) irregular $\hat{u}_i$ and trend $\hat{g}_i$. Following the X11/X12ARIMA procedure, the central 9 point filter is chosen when $I/C < 1$, the 13-term when $1 < I/C < 3.5$, and the 23-term filter for $I/C > 3.5$.

To obtain a LOESS kernel hierarchy, by means of the RKHS methodology, the density f_0 and the corresponding orthonormal polynomials have to be determined.

The probability function is directly derived by $W(t)$ in eq. (16) (Bianconcini, 2006) and is given by

$$f_{0T}(t) = \frac{70}{81}(1 - |t|^3)^3 I_{[-1,1]}(t), \tag{17}$$

where $\frac{70}{81}$ represents the integration constant of the tricube function on the support $[-1, 1]$. f_{0T} is a Beta kernel estimator with parameters $r = 3$ and $s = 3$, that is

$$K(t) = \left(\frac{r}{2B(s+1, \frac{1}{r})} \right) (1 - |t|^r)^s I_{[-1,1]}(t), \tag{18}$$

where $B(a, b) = \int_0^1 t^{a-1}(1 - t)^{b-1} dt$ with $a, b > 0$ is the beta function.

The set of polynomials orthonormal with respect to the tricube density function are not explicitly known, but they can be derived by means of the Gram-Schmidt orthonormalization procedure. Following Brezinski (1980), the LOESS tricube hierarchy up to the third order is here determined using the determinantal expression

$$K_p(t) = \frac{\det(\mathbf{H}^0_{p,1}(t))}{\det(\mathbf{H}^0_p)} f_{0T}(t), \tag{19}$$

where $\mathbf{H}^0_p$ denote the Hankel matrix of order p built from the sequence of moments of f_{0T}, $\{\mu_0, \mu_1, ..., \mu_{2p-2}\}$, and $\det(\mathbf{H}^0_{p,1}(t))$ is the determinant of the matrix obtained by replacing the first column of $\mathbf{H}^0_p$ by $\begin{bmatrix} 1 & t & t^2 & ... & t^{p-1} \end{bmatrix}^T$.

The density function (17) represents the second order kernel in the hierarchy, and the third order tricube kernel, which corresponds to the classical LOESS 2 estimator, is given by

$$\frac{70}{81}(1-|t|^3)^3 \times (\frac{539}{293} - \frac{3719}{638}t^2). \tag{20}$$

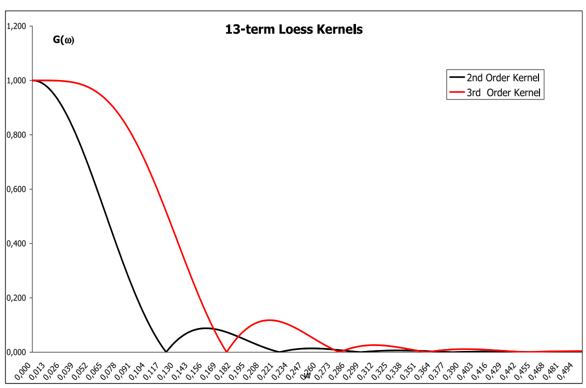

Figure 1: Gain functions of the 13-term symmetric LOESS kernels

Figure 1 illustrates the gain functions of the symmetric 13-term filters in the LOESS tricube hierarchy. Similar results can be extended to every filter length, not shown here for space reasons.

It is evident that the second order kernel possesses the good properties to suppress unwanted ripples (at $w = 0.10$), *i.e.* cycles of 9-10 months often interpreted as false turning points, but at expenses of a too strong reduction of the signal in the output respect to the higher order filter. Hence, the latter is preferable than the former as trend cycle estimator. Particularly, Dagum and Bianconcini (2006a and 2006b) have studied the strong similarity existing between the LOESS 2 estimator and the Henderson filter, that is the most widely applied trend-cycle estimator in nonparametric seasonal adjustment procedures.

Therefore, in the next section, we only derive the asymmetric filters corresponding to the 13-term third order LOESS kernel, and compare them with the classical asymmetric LOESS 2 filters developed by Cleveland (1979).

4 Asymmetric LOESS Smoothers and Their Kernel Representations

The LOESS tricube hierarchy provides a new way to represent the LOESS estimators developed by Cleveland (1979). This has important consequences in the derivation of the corresponding asymmetric filters.

Cleveland (1979) showed that, in the body of the series, LOESS acts as a symmetric moving average with window length $2m+1$. However, at the end of the series, its window length remains $2m + 1$, rather than decreasing to $m+1$ as in the case of the most widely applied asymmetric concurrent trend-cycle estimators. As discussed by Gray and Thomson (1990), this implies a heavier than expected smoothing at the ends of the series respect to the body, and represents a drawback, particularly for economic time series where turning points are important to identify.

On the other hand, in the RKHS approach, the asymmetric weights are directly obtained from the kernel functions adapted to the length of each filter. Given a symmetric filter of length $2m + 1$, we obtain asymmetric smoothers of length ranging from $2m$ to $m + 1$ as follows

$$w_j = \frac{K(j/b)}{\sum_{i=-m}^{q} K(i/b)}, \quad j = -m, ..., q \tag{21}$$

where j is the distance to the target point t_i, b is the bandwidth parameter equal to $m + 1$, and $m + q + 1$ is the asymmetric filter length. For example, the asymmetric weights of the 13-term LOESS kernel for the last point are given by

$$w_j = \frac{K(j/7)}{\sum_{i=-6}^{0} K(i/7))}, \quad j = -6, ..., 0.$$

The asymmetric classical 13-term LOESS 2 filters show an erratic convergence behavior to the corresponding symmetric one, as shown in Figure 2. Furthermore, the last point asymmetric filter, which is the most important for current economic analysis, introduces a strong amplification of the signal at low frequencies and passes a lot of noise.

On the other hand, it is apparent that the asymmetric kernels converge more monotonically and faster to the central one as illustrated in Figure 3. Particularly, the last point LOESS asymmetric kernel exhibits a gain function with better properties of signal passing and noise suppression. This implies smaller filter revisions as new data are added to the series.

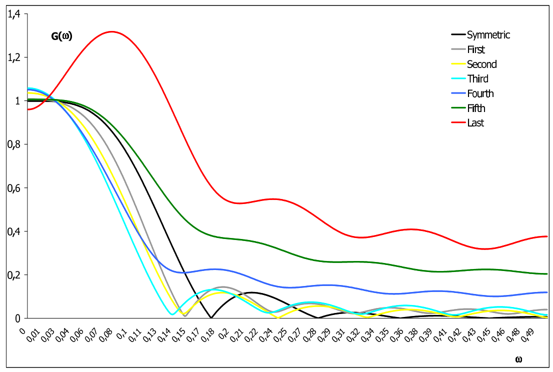

Figure 2: Gain functions of symmetric and asymmetric Classical (Cleveland) 13-term LOESS 2 filters

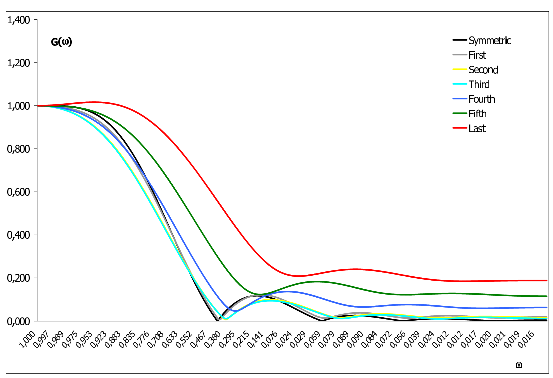

Figure 3: Gain functions of the symmetric and asymmetric weights of the 13-term LOESS Tricube Kernel

The phase shifts for both filters (Figure 4) are smaller than one month in the signal frequency band.

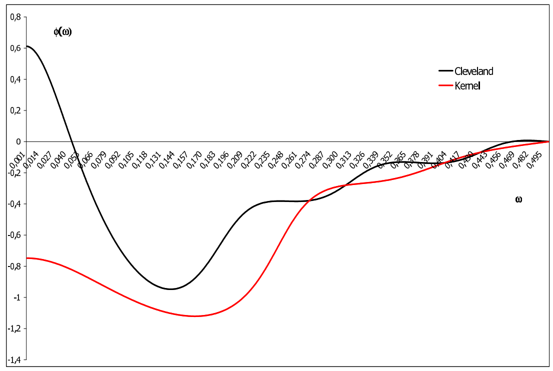

Figure 4: Phase shifts of the asymmetric (end point) weights of the LOESS kernel and of the classical (Cleveland) LOESS 2 filter

5 Empirical Analysis

To illustrate how the "reproducing" kernel estimators respond to the variability of the data we use three series characterized by different noise to signal ratios. The comparison is done as follows:

(1) the input for the trend-cycle estimators is the seasonally adjusted series modified by extreme values with zero weights. The identification and replacement of extreme values is done with the default option of X11ARIMA which defines as extreme value with zero weight any irregular falling outside $\pm 2.5\sigma$;

(2) we obtain preliminary estimates of the trend-cycle by applying the last point asymmetric Cleveland LOESS 2 filter and its kernel representation, and

a final estimate by using the symmetric filter. The length of the smoothers is selected in according to the noise to signal ratio estimated on the series; (3) the comparison between the (last point) asymmetric filters is based on the total revision between preliminary last point estimates and final ones due to only filter change. This is measured by the Mean Absolute Percentage Revision (MAPR) calculated as

$$MAPR = \frac{100}{N} \sum_{t=1}^{N} \left| \frac{\hat{X}_t^P - X_t^F}{X_t^F} \right|, \tag{22}$$

where $\hat{X}_t^P$ is the preliminary trend-cycle estimate at point t obtained by applying the last point asymmetric filter, $\hat{X}_t^F$ is the final trend-cycle estimate at point t obtained by using the symmetric filter, and N is the number of available observations.

The first series is the US Unemployed Men, over 20 years old, covering the period January 1992 - December 1999. This series is adjusted for extreme values as described in (1) above and the Cleveland and kernel LOESS 2 trend-cycle are estimated. We look at the value of the noise to signal ratio given by the X11ARIMA method which is used to select the appropriate length of the filters. The I/C is equal to 0.69 which corresponds to the 9-term filter.

The kernel representation provides more significant gain in filter revisions as exhibit in Figure 5, particularly in the period June 1995 - December 1999, characterized by a larger number of turning points. This pattern is also confirmed by the MAPR equal to 1.20 for the last point asymmetric kernel

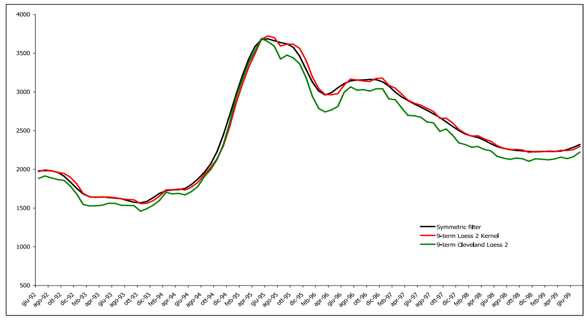

Figure 5: US Unemployed Men (Age over 20) final and preliminary estimates

and to 4.49 for the classical (last point) asymmetric LOESS 2 filter.

The second series is the New Order of Durable Goods (NODG), a Canadian leading indicator discussed in Dagum (1996). This series covers the period January 1981-December 1993, and is published by Statistics Canada in both originally and seasonally adjusted forms. The I/C ratio calculated by the X11/X12ARIMA software is equal to 2.14, and hence we estimate the trend-cycle of the series by means of the last point filters pertaining to the classical 13-term LOESS 2 and its kernel representation as well as the symmetric one.

Figure 6 exhibits the trend-cycle data of both last point asymmetric filters as well as the final one. Clearer than the previous series, the kernel has superior performances in terms of revisions than the classical last point asymmetric filter, with MAPR equal to 1.06 for the former and to 4.00 for the latter.

Finally, the third series is the monthly US Total Exports, covering the period January 1987 - December 1994. The I/C ratio is equal to 4.43 in-

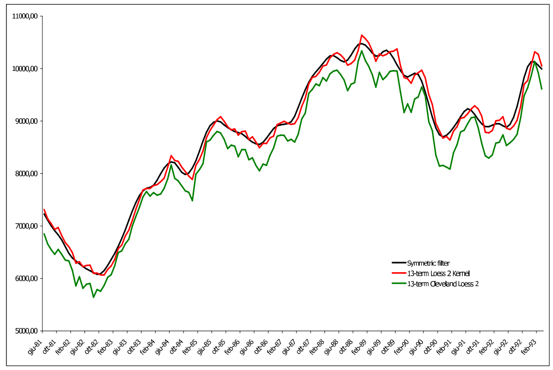

Figure 6: New Order for Durable Goods final and preliminary estimates

dicating that the default option of X11ARIMA will select a 23-term filter as appropriate for trend-cycle estimation. The last point asymmetric kernel performs strongly better than the classical end point filter (Figure 7), with a MAPR equal to 0.88 for the former and equal to 3.14 for the latter. In general, the MAPR tends to decrease when the noise variance is dominant on the signal one, with an increasing better performance for the kernel than the (last point) asymmetric Cleveland filter, that always tends to underestimate the underlying trend.

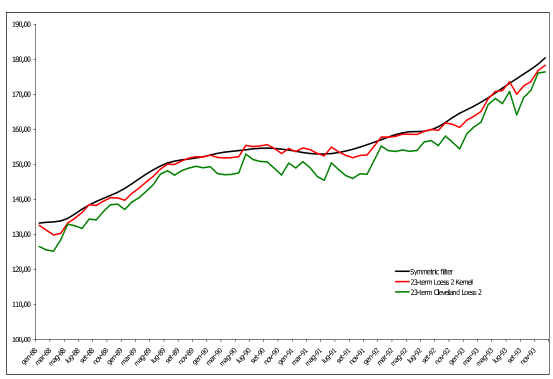

Figure 7: US Total Exports final and preliminary estimates

6 Conclusions

We derived the kernel representation of a locally weighted polynomial smoother which can be applied for current economic analysis, namely the LOESS one. We followed a strictly nonparametric approach of the reproducing kernel Hilbert space methodology by means of a fundamental theorem of Berlinet (1993). This states that a kernel estimator of order $p \geq 2$ can always be decomposed into the product of a reproducing kernel $R_{p-1}(\cdot, \cdot)$ and a probability density function f_0 with finite moments up to order $2p$. The reproducing kernel belongs to the space of polynomials of degree at most $p-1$. An important outcome of the RKHS representation is that the LOESS filters can be grouped into a hierarchy identified by the tricube density f_0, and containing kernels of order 2,3,4,... which are product of orthonormal polynomials by f_0.

Within the LOESS tricube hierarchy, the third order kernel performs better than the lower order one as trend-cycle estimator. In the context of LOESS 2 asymmetric kernel representations it is shown how these possess gain functions with better properties of signal passing and noise suppression relative to the classical ones. Applied to real data characterized by different degrees of noise to signal ratios, for every selected filter length, the reduction in the mean absolute percentage revision (MAPR) is very large for the LOESS last point predictor relative to that developed by Cleveland (1979). The latter always tend to underestimate the underlying trend.

References

[1] Berlinet, A. (1993), Hierarchies of Higher Order Kernels, *Probability Theory and Related Fields*, 94, pp. 489-504.

[2] Bianconcini, S. (2006), *Trend-Cycle Estimation in Reproducing Kernel Hilbert Spaces*, Ph.D. Thesis, Department of Statistics, University of Bologna.

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

[4] Cleveland, W.S.(1979), Robust Locally Regression and Smoothing Scatterplots, *JASA*, 74, 829-836.

[5] Cleveland, W.S. and Devlin, S.J. (1988), Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting, *JASA*, 83, 596-610.

[6] Cleveland, R., Cleveland, W., McRae, J. and Terpenning, I. (1990), STL: A Seasonal Trend Decomposition Procedure Based on LOESS, *Journal of Official Statistics*, 6, 3-33.

[7] Dagum, E. B. (1985), Moving Averages, in *Encyclopedia of Statistical Sciences*, S. Kotz and N. Johnson, editors, John Wiley and Sons, 5, pp. 631-634.

[8] 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.

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

[10] Dagum, E.B. and Bianconcini, S. (2006a), The Henderson Smoother in Reproducing Kernel Hilbert Space, *Eurostat Conference on "Seasonality, Seasonal Adjustment and Their Implications for Short-Term Analysis and Forecasting*, May 10-12, 2006.

[11] Dagum, E.B. and Bianconcini, S. (2006b), Local Polynomial Trend-Cycle Predictors in Reproducing Kernel Hilbert Spaces for Current Economic Analysis, *Anales de Economia Aplicada*, pp. 1-22

[12] Dagum, E.B. and Luati, A. (2000), Predictive Performances of Some Linear and Nonlinear Smoothers for Noisy Data, *Statistica*, 4, 635-654.

[13] Dagum, E.B. and Luati, A. (2004), Relationship Between Local and Global Nonparametric Estimators Measures of Fitting and Smoothing, *Studies in Nonlinear Dynamics and Econometrics*, 8, article 17.

[14] Fan, J. and Gijbels, I. (1996), *Local Polynomial Modelling and Its Applications*, Chapman and Hall, New York.

[15] Findley, D., Monsell, B, Bell, W., Otto, M. and Chen, B. (1998), New Capabilities and Methods of the X12ARIMA Seasonal Adjustment Program, *Journal of Business Economic Statistics*, 16, pp. 127-152.

[16] Gray, A. and Thomson, P. (1990), Comments on STL: A Seasonal Trend Decomposition Procedure Based on LOESS, *Journal of Official Statistics*, 6, pp.47-55.

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

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

[19] Kendall, M.G., Stuart, A. and Ord, J. (1983), *The Advanced Theory of Statistics*, Vol. 3, Ed. C. Griffin.

[20] 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.

[21] Loader, C. (1999), *Local Regression and Likelihood*, Springer, New York.

[22] Macauley, F.R. (1931), *The Smoothing of Time Series*, National Bureau of Economic Research, New York.

[23] Mallows, C.L. (1966), Choosing a Subset Regression, *Annual Meeting American Statistical Association*, Aug 15-19.

[24] Mallows, C.L. (1966), Some Comments on C_p, *Technometrics*, 15, pp. 661-675.

[25] Moore, G.H. (1961), *Business Cycle Indicators*, Princeton University Press, Princeton, New Jersey, US.

[26] Shiskin, J., Young, A. and Musgrave, J. (1967), *The X-11 Variant of the Census Method II Seasonal Adjustment Program*, Technical Paper 15, US Department of Commerce, Bureau of the Census, Washington DC.

[27] Wand, M.P. and Jones, M.C. (1995), *Kernel Smoothing*, Monographs on Statistics and Applied Probability, 60, Chapman and Hall.

[28] Whittaker, E.T. and Robinson, G. (1924), *Calculus of Observations: a Treasure on Numerical Calculations*, Blackie and Son, London.