Econometric Reviews
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/lecr20

A Review of Some Modern Approaches to the Problem of Trend Extraction

Theodore Alexandrov [a], Silvia Bianconcini [b], Estela Bee Dagum [b], Peter Maass [a] & Tucker S. McElroy [c]

[a] Center for Industrial Mathematics, University of Bremen, Bibliothekstr. 1, Bremen, Germany

[b] Department of Statistics, University of Bologna, Bologna, Italy

[c] Statistical Research Division, U.S. Census Bureau, Washington, D.C., USA

Available online: 17 Oct 2011

To cite this article: Theodore Alexandrov, Silvia Bianconcini, Estela Bee Dagum, Peter Maass & Tucker S. McElroy (2012): A Review of Some Modern Approaches to the Problem of Trend Extraction, Econometric Reviews, 31:6, 593-624

To link to this article: http://dx.doi.org/10.1080/07474938.2011.608032

Econometric Reviews, 31(6):593–624, 2012

ISSN: 0747-4938 print/1532-4168 online
DOI: 10.1080/07474938.2011.608032

A REVIEW OF SOME MODERN APPROACHES TO THE PROBLEM OF TREND EXTRACTION

Theodore Alexandrov[1], Silvia Bianconcini[2], Estela Bee Dagum[2], Peter Maass[1], and Tucker S. McElroy[3]

[1] *Center for Industrial Mathematics, University of Bremen, Bibliothekstr. 1, Bremen, Germany*
[2] *Department of Statistics, University of Bologna, Bologna, Italy*
[3] *Statistical Research Division, U.S. Census Bureau, Washington, D.C., USA*

□ *This article presents a review of some modern approaches to trend extraction for one-dimensional time series, which is one of the major tasks of time series analysis. The trend of a time series is usually defined as a smooth additive component which contains information about the time series global change, and we discuss this and other definitions of the trend. We do not aim to review all the novel approaches, but rather to observe the problem from different viewpoints and from different areas of expertise. The article contributes to understanding the concept of a trend and the problem of its extraction. We present an overview of advantages and disadvantages of the approaches under consideration, which are: the model-based approach (MBA), nonparametric linear filtering, singular spectrum analysis (SSA), and wavelets. The MBA assumes the specification of a stochastic time series model, which is usually either an autoregressive integrated moving average (ARIMA) model or a state space model. The nonparametric filtering methods do not require specification of model and are popular because of their simplicity in application. We discuss the Henderson, LOESS, and Hodrick–Prescott filters and their versions derived by exploiting the Reproducing Kernel Hilbert Space methodology. In addition to these prominent approaches, we consider SSA and wavelet methods. SSA is widespread in the geosciences; its algorithm is similar to that of principal components analysis, but SSA is applied to time series. Wavelet methods are the de facto standard for denoising in signal procession, and recent works revealed their potential in trend analysis.*

Keywords Model-based approach; Nonparametric linear filtering; Singular spectrum analysis; Time series; Trend; Wavelets.

JEL Classification C01; C02; C14; C40; C50.

Disclaimer. This report is released to inform interested parties of research and to encourage discussion. Any views expressed on statistical issues are those of the authors and not necessarily those of the U.S. Census Bureau or the Universities of Bremen and Bologna.

Address correspondence to Theodore Alexandrov, Center for Industrial Mathematics, University of Bremen, Bibliothekstr. 1, Bremen 28359, Germany; E-mail: theodore@math.uni-bremen.de

1. INTRODUCTION

The identification and estimation of trends is of vital importance to the field of econometrics because of the wide interest in studying the long terms movements of economic time series. Mathematical approaches to trend extraction have a long history, going back to the early days of statistics with the advent of least-squares regression. In the nineteenth century trend extraction consisted of fitting a simple function (usually a linear one) to the data. With the rapid growth of the theory of stationary time series the trend was conceived of as a non-stochastic component, which would be subtracted out in order to obtain a stationary time series that afterwards could be successfully modeled (see Wold, 1938). At the present time stochastic approaches to the definition of a trend are widely used in econometrics, since it is popular to view an economic time series as the aggregation of a myriad of individual factors. In this article, we present the contemporary research on the trend extraction problem for one-dimensional economic time series, considered from a variety of perspectives.

We define a *time series* of length N as $X = (x_0, \ldots, x_{N-1})$, $x_n \in \mathbb{R}$. There are a variety of definitions of trend, but all of them imply the following additive model:

$$x_n = t_n + r_n, \qquad \text{or} \qquad X = T + R, \tag{1}$$

where $T = (t_0, \ldots, t_{N-1})$ denotes a trend and $R = (r_0, \ldots, r_{N-1})$ is referred to as the *residual*. The latter can have both predictable (i.e., it has one-step ahead forecast errors equal to zero) and non-predictable components d_n and s_n. Hence, we come to the expansion

$$x_n = t_n + d_n + s_n. \tag{2}$$

Model (2) with a periodic d_n and a zero mean s_n is referred to as the *classical decomposition model* (Brockwell and Davis, 2003).

In his remarkable book, Chatfield (1996) defines trend as "a long-term change in the mean level," and expresses a typical viewpoint about trends by considering them as smooth additive components containing the information about global change in the given time series. The problem of the extraction of such a component occurs in many applied sciences, and attracts scientists from diverse areas that apply their knowledge for its solution. As a result, there exist several different methods for trend extraction, some of which have been developed by or in collaboration with econometricians. These methods differ in their complexity and interpretability, as well as the mathematical tools that they use, and hence produce dissimilar results. It is difficult to assess whether a given method is "best," because the actual criteria are vague and subject to controversy; moreover, each method may be "best" within its intended context, when the assumptions undergirding its construction are satisfied, but may well perform worse on time series that greatly deviate from those core assumptions. In practice, the experts in each area of application will decide which features of trends are most important for their research goals, and will choose trend extraction methods that are best-suited to reproduce those features.

This article focuses on describing and contrasting the most prevalent trend estimation methods that are currently used in econometrics: model-based trend extraction and nonparametric filtering, as well as two promising novel approaches to trend extraction, namely, singular spectrum analysis (SSA), and wavelets. The model-based approach (MBA) utilizes a specification of a stochastic time series model that is typically either an autoregressive integrated moving average (ARIMA) model or a state space model. The developments of such methods were heavily influenced by trend extraction problems in the engineering and econometric disciplines. Currently, they are extremely popular in econometrics, and are also used in many other scientific areas. The *nonparametric filtering* methods do not require specification of a model; they are quite easy in application and are used in all applied areas of time series analysis. The algorithm of SSA is identical to that of principal components analysis (up to normalization of the covariance matrix), but SSA is instead applied to time series. It mainly originated in dynamical systems and at the present time is widespread in the geosciences, but has recently been used in econometric studies as well. Wavelets-based methods are currently a de facto standard for denoising in many scientific applications, and have been used to produce trends from financial and macro-economic time series. We provide an overview of all these methods, and how they are used to extract trends from economic data.

Of course, there are many other methods for trend extraction (polynomial fitting, demographic methods, Bayesian estimation, etc.), but the majority of econometric methods follow one of the four approaches considered in this article. The rest of the article is organized as follows. In Section 2 we discuss the problem of trend extraction, setting out basic terminology and notation common to all methods. Section 3 reviews the MBA, tracing the history of the subject through the econometric, statistical, and engineering literature up to the present time; we then discuss the state-of-the-art of MBA trend extraction by delineating the main competing methodologies. In Section 4 we describe the nonparametric linear filtering methods based upon the Henderson, the LOcally wEighted Scatterplot Smoothing (LOESS), and the Hodrick–Prescott filters. Then we briefly discuss how using the Reproducing Kernel Hilbert Space methodology improves their properties. In Section 5 we consider SSA, tracing its history from the literature on dynamical systems up through the present time, as it is being increasingly implemented in econometrics. Wavelet methods for trend extraction are reviewed in Section 6, with a focus on their foundations in the fields of mathematics and statistics. Finally, in Section 7 we present a table summarizing characteristics of the considered approaches and a demonstration of the application of the methods to a real time series (electric power) is examined. Even though one cannot derive general conclusions from this single example, nevertheless it serves to illustrate the most salient and vital aspects of each of the methods.

2. THE TREND EXTRACTION PROBLEM

For a time series satisfying model (1), the *problem of trend extraction* is defined as estimation of an unknown T having only X. Note that in the literature devoted to state space methods (see Section 3) one refers to the *smoothing problem* as estimating the whole vector T from X (Durbin and Koopman, 2001). However, if we only desire t_{N-1}, i.e., the most current value of the trend, then this is called the *filtering problem*.

2.1. Non-Stochastic and Stochastic Trend Definition

Following the *non-stochastic approach*, the trend is defined as a non-stochastic function belonging to some class. The function class can be given explicitly, specifying a parametric model, or implicitly, imposing some smoothness condition. A widespread example of the parametric trend is a polynomial trend and its elementary version, a linear trend. Due to the simplicity of its estimation (e.g., using least-squares regression), removal of a polynomial trend is used even in stochastic time series modeling, e.g., see examples in (Pollock, 2008). The smoothness condition is usually formulated in terms of derivatives or Fourier coefficients.

Another approach to the definition of trend is the *stochastic approach*, which defines trend as a realization of a stochastic process. The trend is again supposed to be smooth where the smoothness is expressed in terms of the variance or autocorrelation function (Froeb and Koyak, 1994). An elementary example of a stochastic trend is a random walk with a drift. In the stochastic approach, one usually assumes the orthogonality between the trend and the residual (the latter is generally supposed to be stochastic); for more details see Section 3. Note that the stochastic approach can model the non-stochastic trends considered above. For example, a polynomial of degree $d-1$ is a component of an integrated process of d unit roots defined over a finite interval.

Note that economists and statisticians are often interested in the "short" term trend of socio-economic time series. The short term trend generally includes cyclical fluctuations, and is referred to as *trend-cycle*.

The MBA considers stochastic trends, whereas both nonparametric filtering and wavelet methods assume a globally smooth trend, such that it can be locally approximated by a non-stochastic function of time. At the same time, some nonparametric filters (e.g., the Hodrick–Prescott filter) can be derived based on stochastic principles; moreover, their properties can be studied in the stochastic framework, see Section 4 for more details. SSA resembles Karhunen–Loéve decomposition of stochastic processes but with a large part of its analysis formulated without reference to stochastic theory; see Section 5.

2.2. Updating Trend with New Data

In signal processing, the filter which uses only past and present values is called a *causal filter*. In econometrics, such a filter is named a *concurrent filter*.

Given a time series where more data will arrive in the future, one has two options for presenting trends: the online approach and the window approach. The *online approach* generally uses concurrent filters (or trend extraction methods that do not require future data), and each time a new data point arrives, one generates the next concurrent estimate. Such a method does not involve trend revisions, since once a concurrent filter is used to produce a trend value at a given time point, it is not revisited later, see Wildi (2005) for a discussion for the motivations behind online estimates of trend. Exponential smoothing is a popular example of online filtering (Durbin and Koopman, 2001).

The *window approach* produces trend estimates at every time point in the available sample, which is accomplished either through the use of finite-length filters or through forecast-extension of the series. When new data arrives, trend estimates in the middle of the sample must be updated, or revised. The difference between a new trend estimate (based upon more current data) and an old trend estimate is called a revision; for references on revisions in the context of seasonal adjustment, see Pierce (1980), Findley et al. (1998), and Maravall and Caporello (2004).

2.3. Difference from Denoising Problem

If the residual after trend extraction corresponds to noise, then the problem of trend extraction is similar to the problem of *denoising*. However, denoising methods are usually constructed taking into account only the noise distribution, whereas any trend extraction procedure considers the following trend properties. First, the trend is assumed to be smooth, in contrast to the denoised signal. Second, the trend is to be separated from a seasonal component and other evident periodic components. Moreover, in denoising the residual is usually assumed to follow one of the typical noise models (i.e., white or red noise), whereas in trend extraction the residual may follow much more general models.

2.4. Trend Detection

Another trend-related problem is to determine whether the given time series contains a trend significantly different from zero. This problem often occurs in geosciences (e.g., establishing the existence of an increasing trend in temperatures) and is usually referred to as the *trend detection*. For the solution, the statistical tests are used which require the specification of the trend model.

The detection of monotonic trends is the most extensively studied. One of the widespread tests for this problem is the non-parametric *Kendall* (or *Mann–Kendall*) *test* and its modifications for handling seasonality (Hirsch and Slack, 1984) and autocorrelated data (Hamed and Rao, 1998). Among others are: the parametric *t-test* and the nonparametric *Mann–Whitney* and *Spearman tests*. Berryman et al. (1988) describes many methods for monotonic trend detection and provides a selection algorithm for them. For more recent developments see the review of Esterby (1996) which is focused on hydrological applications; also see a comparative study (Yue and Pilon, 2004) describing several bootstrap-based tests.

3. MBA

3.1. Preamble

The MBA to trend estimation refers to a family of methods, which have in common the reliance upon time series models for the observed, trend, and residual processes. The history of this approach is briefly discussed below.

In the discussion of trend estimation in this article, we naturally focus upon finite samples, since this is the only data available in practice. The early literature on MBA signal extraction developed the theory for doubly-infinite and semi-infinite samples (Whittle, 1963; Wiener, 1949), and exclusively focused on stationary processes. We also note that this early theory encompassed continuous-time processes as well, since early "filters" were essentially given by the operation of analog-type hardware; however we do not pursue continuous-time trend estimation here; see Koopmans (1974) for a discussion.

The MBA literature on trend extraction began to be generalized in two directions: dealing with boundary effects (i.e., the finite sample) and handling nonstationarity (generally speaking, homogeneous nonstationarity exemplified by ARIMA processes). The engineering community focused on the former, the pivotal discovery being the so-called

Kalman filter (Kalman, 1960). Rauch (1963) extended the Kalman filter to handle boundary effects; these algorithms rely on a state space formulation (SSF) of trend extraction. Additional discussion of state space methods from an engineering perspective can be found in Anderson and Moore (1979) and Young (1984).

However, since engineers are primarily concerned with stationary data, the SSF approach did not handle nonstationary data until econometricians became involved later on. Books that discuss SSF from an econometrics/statistical perspective include Harvey (1989), Kitagawa and Gersch (1996), and Durbin and Koopman (2001). Generally speaking, trends are nonstationary processes, so the basic stationary approach of the older engineering literature is not adequate.

In our description of MBA, we focus upon methods following the window approach (see Section 2.2), although the discussion is easily adapted to include online-trend extraction, i.e., concurrent filters.

General Questions

The MBA method of trend estimation generally requires a specification of the dynamics of signal and noise (trend and residual). These are typically considered to be stochastic, but possibly include a component that is completely predictable from its infinite past. This predictable portion typically takes the form of a polynomial or sinusoidal function with random coefficients, and will be referred to as "analytic"; see Slepian (1976). In order to specify signal and noise dynamics, the MBA method requires some thought about the following issues, in more or less this order: (i) How are the trend and residual processes related to the observed process? (ii) How are trend and residual related to one another? and (iii) How are trend estimates to be generated? and (iv) What types of models are being considered, and how are they estimated?

3.2. Trend in MBA

Some approaches to model-based trend estimation view the trend as a direct function of the data, e.g., the trend is a subjective summary of the low-frequency features of the data. This assumption is implicitly understood in many of the trend estimation approaches in the engineering literature; also see the Direct Filter Approach of Wildi (2005). In contrast, the trend may be viewed as an objective entity. This is a more common approach among statisticians, and generally entails the development of a model for T.

Relation between Trend and Residual

The two most popular assumptions which regulate the relations between trend and residual are the *orthogonal* and *Beveridge–Nelson* (Beveridge and Nelson, 1981), or BN. The BN assumes that both trend and residual can be written as linear filters of the data innovation process, and thus are "fully" correlated. The *orthogonal approach* assumes that "differenced" trend and residual (i.e., the components after nonstationary effects have been removed by differencing) are uncorrelated with one another. This supposition is more consistent with economic data, since diverse components are thought to originate from diverse aspects of the economy, and thus should not be correlated. Naturally, the orthogonal decomposition and BN decomposition represent the opposite ends of the spectrum; some work by Proietti (2006) deals with the case that the components are less than fully correlated.

3.3. Construction of the Trend Model

For MBA trend extraction, we require a model for the trend and residual; note that this residual may contain seasonal effects, and can therefore be nonstationary. There are several popular approaches for obtaining these models from the data: Decomposition, Structural, and BN.

The *Decomposition approach* (Burman, 1980; Hillmer and Tiao, 1982) begins by attempting to fit an optimal model to the observed data, where optimality is often equated with maximum likelihood estimation of model parameters after different model specifications have been compared via information criteria (e.g., Akaike Information Criterion or other goodness-of-fit diagnostic tests; see Findley et al., 1998, for a discussion). Then the models for trend and residual are determined via partial fraction decomposition techniques applied to the autocovariance generating function of the model for the data (this assumes that ARIMA or Seasonal ARIMA models are being used). Some amount of user-specification is required, since all differencing operators and autoregressive operators that appeared in the data model must be allocated (subjectively) to the various components. For more discussion of this, see Bell and Hillmer (1984) and Chapter 8 of Peña et al. (2001). Typically there is indeterminancy of the derived component models; maximizing the variance of the irregular component results in the *Canonical Decomposition*, which results in signals that are as stable as possible.

The *Structural approach* (Harvey, 1989) on the other hand also uses maximum likelihood estimation, but the form of the likelihood is dictated by a pre-specified model form for the components. This is also referred to as an Unobserved Components (UC) approach. While sometimes a canonical decomposition does not exist (see Hillmer and Tiao, 1982), the structural approach is always viable. However, the implied model for the data process will be a parameter-restricted ARIMA model, in contrast to the (unconstrained) ARIMA model of the decomposition approach. Also, more a priori information about the trend and residual dynamics are needed from the user, such as specifying the differencing order for the trend ahead of time. See Durbin and Koopman (2001) for more discussion.

Note that the term *Structural Model* refers to a particularly simple class of component models promoted by Gersch and Kitagawa (1983), which are essentially parameter-restricted ARIMA models. Here we distinguish between the Structural Approach to estimating component models (which can be general ARIMA models) and the more specific Structural Models utilized in STAMP and SsfPack (Koopman et al., 1999).

The *BN approach* is much like the Structural approach, though now the component models are fully correlated; it is also based on an ARIMA model of the data, in common with the *Decomposition approach*. However, in contrast with the latter, the *BN approach* utilizes a partial fraction decomposition of the ARIMA transfer function, not the autocovariance generating function. Although the likelihood takes a different form, since the components are not orthogonal, we can still utilize maximum likelihood estimation to get the component models, and the corresponding trend filters are then easy to obtain; see Proietti (2006).

3.4. Penalty Function

The next issue is: What sort of penalty function is used to determine optimal signal extraction? Mean squared error (MSE) is very popular among statisticians, and is the original penalty function used in Wiener (1949). The conditional expectation of the trend T given the data X is the estimate that minimizes MSE; if in addition the data has a Gaussian distribution, the trend estimate will be linear in the data. This explains the central role of linear estimators in the MBA literature. However, when alternative distributions are present (e.g., log-normals), other penalty functions such as Relative MSE may be more appropriate; see Thomson and Ozaki (2002).

If we are interested in linear estimators, how do we find the best one? Bell (1984) discusses how these are computed from the autocovariance generating functions of trend and residual, assuming certain conditions on the initial values of the data; also see Cleveland and Tiao (1976). Adaptations of this theory to finite samples can be found in Bell (2004), Pollock (2007), and McElroy (2008). The following references discuss the finite-sample theory from an SSF viewpoint: Koopman (1997) and Durbin and Koopman (2001). An excellent discussion that contrasts the SSF approach with non-model-based methods can be found in Young and Pedregal (1999).

Given that most statistical readers will be interested in linear MSE-optimal estimators, we focus on the SSF approach and the matrix approach, which are equivalent. The widely-used SSF technique requires *Assumption A* of Bell (1984) for its estimates to be MSE-optimal, as is discussed in McElroy (2008). Assumption A states that the d initial values of the process x_n are independent of differenced trend and differenced residual. Note that the SSF estimator is a linear operation on the data X that produces a vector of trend estimates T; the linear matrix that accomplishes this is derived in McElroy (2008). For some purposes, it is convenient to have this matrix, e.g., the full error covariance matrix is easily obtained using this approach.

3.5. Model Classes

Finally, the MBA requires a choice of model classes. As mentioned above, the Decomposition approach relies on seasonal ARIMA models for the components. The ARIMA and Structural models (Harvey, 1989) are the most popular models in econometric MBA trend estimation. In theory, one only needs the autocovariance generating function for trend and residual in order to proceed. For example, another class of models are time-varying coefficient models, where the parameters evolve according to a random walk or other such process; a discussion of using such models can be found in Young et al. (1999).

3.6. Software

Several of the main software packages for MBA trend estimation include X-12-ARIMA (Findley et al., 1998), TRAMO-SEATS (Maravall and Caporello, 2004), STAMP (Koopman et al., 2000), and microCAPTAIN (Young and Benner, 1991). X-12-ARIMA mixes nonparametric linear filtering with model-based forecast and backcast extension, so it can be viewed as a partial MBA. SEATS is fully MBA, and utilizes the Canonical Decomposition approach. On the other hand, STAMP utilizes a Structural Approach as well as Structural Models. The program microCAPTAIN also uses a Structural Approach, but with time-varying coefficient models that are estimated using the frequency domain method of Dynamic Harmonic Regression (DHR), as opposed to the Maximum Likelihood Estimation method of the other software (Young et al., 1999). These are some of the core MBA software products for trend estimation, of which some other products (DEMETRA of EuroStat, SAS implementations of Structural Models) are derivatives. We also mention the RegComponent software (Bell, 2004), which uses a Structural Approach with ARIMA models (although extensions now allow for a Decomposition Approach as well);

this was one of the first programs to simultaneously handle smoothing by SSF methods and also estimate fixed regression effects.

4. NONPARAMETRIC TREND PREDICTORS

4.1. Preamble

In the nonparametric trend filtering approach the model (2) is usually considered, where the time series X is supposed to be seasonally adjusted, T is referred to as a trend-cycle, and R is assumed to be either a white noise, $NID(0, \sigma^2)$, or, more generally, to follow a stationary and invertible Autoregressive Moving Average process. Expecting that the trend-cycle T is smooth, it can be locally approximated by a d-degree polynomial. The method of local polynomial regression entails the fitting of a succession of polynomials to the points that fall within a moving window of a fixed width. The polynomial that is estimated from the $q = 2m + 1$ points $x_n + j$, $j = -m, \ldots, m$, is denoted by

$$t_n(j) = a_0 + a_1 j + \cdots + a_d j^d, \tag{3}$$

where $a_k \in \mathbb{R}$, $k = 0, \ldots, d$. Here, n is the index on the central point x_n, for which the polynomial provides a smoothed value in the form of $t = t_n(0) = a_0$, which is the estimated trend value that replaces x_n. It can be shown that the method of local polynomial regression amounts to the application of a symmetric moving average filter to the points that fall within the window

$$\hat{t}_n = \sum_{j=-m}^{m} b_j x_{n-j}. \tag{4}$$

The applied weights $\{b_j\}_{j=-m}^{m}$ depend on: (i) the degree d of the fitted polynomial, (ii) the filter span $2m + 1$, and (iii) the shape of the function used to average the observations in each neighborhood.

The local polynomial regression predictor developed by Henderson (1916) and LOESS due to Cleveland (1979) are the most widely applied nonparametric local filtering methods to estimate the short-term trend of seasonally adjusted economic indicators. In this section we also consider the Hodrick and Prescott (1997) filter which is widely used for economic and financial applications.

4.2. Henderson Filter

The Henderson filters are derived from the graduation theory, known to minimize smoothing with respect to a third degree polynomial within the span of the filter. The minimization problem

$$\min_{a_k, 0 \le k \le 3} \left\{ \sum_{j=-m}^{m} w_j \left[x_{t+j} - a_0 - a_1 j - a_2 j^2 - a_3 j^3 \right]^2 \right\} \tag{5}$$

is considered, where the symmetric weights w_j are chosen to minimize the sum of squares of their third differences (*smoothing criterion*). This filter has the property that fitted to exact cubic functions will reproduce their values, and fitted to cubic polynomials affected by additive noise, it will give smoother results than those obtained by ordinary least squares. Henderson (1916) proved that two alternative smoothing criteria give the same formula, as shown explicitly by Kenny and Durbin (1982): (i) minimization of the variance of the third differences of the series $\hat{t}_n$ defined by the application of the moving average in (4) and (ii) minimization of the sum of squares of the third differences of the coefficients b_j of the moving average formula (4). Moreover, Henderson (1916) showed that the nth element of the trend estimation, $\hat{t}_n$, is given by

$$\hat{t}_n = \sum_{j=-m}^{m} \phi(j) w_j x_{n-j},$$

where $\phi(j)$ is a cubic polynomial whose coefficients have the property that the smoother reproduces the data if they follow a cubic. Henderson also proved the converse: if the coefficients b_j of a cubic-reproducing summation formula 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 $\phi(j) w_j = b_j$. To obtain w_j from b_j one simply divides b_j by a cubic polynomial whose roots match those of b_j.

The asymmetric filters commonly used in association with the Henderson smoother were developed by Musgrave (1964) on the basis of minimizing the mean squared revision between final and preliminary estimates. Although the basic assumption is that of fitting a linear trend within the span of the filter, the asymmetric weights can only reproduce a constant for the only imposed constraint is that the weights add to one. Important studies related to these kind of trend-cycle estimators have been made, among many others, by Kenny and Durbin (1982), and Dagum and Bianconcini (2008).

4.3. LOESS Filter

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. This filter consists of locally fitting polynomials of degree d by means of weighted least squares on a neighborhood of q observations around the estimated point. As q increases, the estimated trend $\hat{t}_n$ becomes smoother. Let $x_{n-m}, \ldots, x_{n+m}$ be $q = 2m + 1$ points in the neighborhood of x_n to which the polynomial function $t_n(j)$ of degree d is fitted by a weighted least square regression that minimizes the function

$$\sum_{j=-m}^{m} w_j \left[x_{n+j} - \sum_{k=0}^{d} a_k j^k \right]^2 .$$

The LOESS estimator is similar to the Henderson one, but the former (i) allows one to fit polynomials of degree d, (ii) is defined everywhere, and (iii) does not impose the Henderson smoothing criterion on the weights. The degree d of the fitted polynomial, the shape of the weight function w, and the value of the smoothing parameter q are the three crucial choices to be made in LOESS.

Polynomials of degree $d = 1$ or $d = 2$ are generally suitable choices. On the assumption that the points are equally spaced, the weighting function of the LOESS trend estimator can be given by the tricube function $w(j) = (1 - |j/(m+1)|^3)^3$ as proposed by Cleveland (1979). It will be observed that this function is symmetric about the point $j = 0$ and it reaches a minimum when $j = \pm m$ (for $|j| = m + 1$, it would be zero-valued). The Loess procedure also applies a discount to the weights associated with locally aberrant observations, which are outliers in other words. Let $r_j = x_j - t_j$ be the deviation of a data point from the current fitted value. Then the proposed discount factor is $d_j = (1 - \{r_j/6s\}^2)^2$, where s is the median value of the deviations amongst the local set of q data points. If the discount factor becomes negative, then it is replaced by a zero. The discount factor can be used in a second round of estimation in which the original weights w_j are replaced by the discounted weights $d_j w_j$, as done, for example, in a Seasonal-Trend decomposition procedure based on LOESS (STL) developed by Cleveland et al. (1990).

The ratio between the amplitude of the neighborhood q and the full span of the series N defines the *smoothing parameter*. It is sensible to choose an odd value for q in order to allow symmetric neighborhoods for central observations. A low smoothing parameter gives unbiased (in the sense of reproducing a polynomial trend of degree d without distorsion) but highly variable estimates, while increasing its value reduces the variance but augments the bias. In choosing the smoothing parameter, the aim is to take a large q in order to minimize the variability in the smoothed points but without distorting the underlying trend.

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. However, this implies a heavier than expected smoothing at the ends of the series with respect to the body, and represents a drawback, particularly for economic time series, where it is important to identify turning points.

4.4. Hodrick–Prescott (HP) Filter

In Hodrick and Prescott (1997), the authors derived a filter in the manner pursued by Reinsch (1967) in deriving the closely related cubic smoothing spline. The framework used in Hodrick and Prescott (1997) is that a given time series x_n is the sum of a growth component t_n and a cyclical component c_n: $x_n = t_n + c_n$. The *measure of the smoothness* of the trend t_n is the sum of the squares of its second difference. The c_n are deviations from t_n and the conceptual framework is that over long time periods, their average is near zero. These considerations lead to the following programming problem of estimation of the trend t_n

$$\{\hat{t}_n\}_{n=0}^{N} = \arg \min_{\{t_n\}_{n=0}^{N-1}} \left\{ \sum_{n=0}^{N-1} (x_n - t_n)^2 + \lambda \sum_{n=0}^{N-1} [(t_n - t_{n-1}) - (t_{n-1} - t_{n-2})]^2 \right\}. \tag{6}$$

The parameter λ is a positive number which penalizes variability in the growth component series. The larger the value of λ, the smoother the solution series. The limit of the solution to (6) as λ approaches infinity is the least squares fit of a linear time trend model. For economic and financial quarterly series, lambda is recommended to be equal to 1600.

The Hodrick–Prescott (HP) filter was not developed to be appropriate, much less optimal, for specific time series generating processes. Rather, apart from the choice of λ, the same filter is intended to be applied to all series. Nevertheless, the smoother that results from the solution of Eq. (6) can be viewed in terms of optimal signal extraction literature pioneered by Wiener (1949) and extended by Bell (1984) to incorporate integrated time series generating processes. King and Rebelo (1993) analyzed the HP filter in this framework, motivating it as a generalization of the exponential smoothing filter. On the other hand, it can be shown that, according the Wiener–Kolmogorov principle and on the assumption that the distributions are Gaussian, the HP filter would provide the optimal estimate to the trajectory of an integrated random walk when it has been obscured by errors of observation that are independently and identically distributed. Such an error-contaminated process can be described by an $IMA(2, 1)$ model, i.e., an $ARIMA(0, 2, 1)$ model (see, e.g., Kaiser and Maravall, 2001). On the other hand, Harvey and Jaeger (1993) interpreted the HP filter in terms of Structural Models (see Section 3.5). Several authors have analyzed shortcomings and drawbacks of the filter, concentrating on the stochastic properties of the estimated components induced by the filter. We refer to Ravn and Uhlig (1997) for a detailed summary.

4.5. Filters in Reproducing Kernel Hilbert Space

The Henderson and LOESS filters, aimed at extracting local polynomial trends, can be derived using the *Reproducing Kernel Hilbert Space* (RKHS) methodology. Moreover, this approach leads to new nonparametric estimators outperforming original Henderson and LOESS filters.

The theory and systematic development of reproducing kernels and associated Hilbert spaces was laid out by Aronszajn (1950). Recently, reproducing kernel methods were used as a framework for penalized spline methodology (Wahba, 1990) and other applications. For the lack of space, we do not explain the notion of RKHS in this article but refer to Dagum and Bianconcini (2006). The main idea of deriving a filter using the RKHS methodology is as follows.

Let us introduce a Hilbert space $L^2(f_0)$ of functions with the inner product $\langle u, v \rangle = \int_{\mathbb{R}} u(t)v(t)f_0(t)dt$, where f_0 is a probabilistic density function weighting each observation to take into account its position in time; f_0 plays the same role as the weights w_j used in (5). Let $\mathbf{P}_d$ be a Hilbert subspace of $L^2(f_0)$ containing polynomials of degree at most d (d is a non-negative integer). Then, fitting a local polynomial trend to data x_n is equivalent to projecting the data x_n onto $\mathbf{P}_d$.

Berlinet (1993) showed that the space $\mathbf{P}_d$ is a reproducing kernel Hilbert space. Moreover, using RKHS properties he proved that the coefficients b_k of the filter (4) can be calculated as products of the reproducing kernel of $\mathbf{P}_d$ and a density function f_0.

Hence, one can derive a filter using this approach having (i) a reproducing kernel of $\mathbf{P}_d$, and (ii) an appropriate density f_0. Naturally, the density depends on the desired span width of the filter and the kernel depends on the degree d of the fitted polynomial. Dagum and Bianconcini (2006) and Dagum and Bianconcini (2008) have found reproducing kernels in Hilbert spaces of the Henderson and LOESS local polynomial regression predictors with particular emphasis on the asymmetric filters applied to the most recent observations. The RKHS-based variants of the Henderson and LOESS filters are shown to have superior properties relative to the classical ones from the view point of signal passing, noise suppression and revisions.

What is more, an important outcome of the RKHS theory is that the resulting linear filters can be grouped into hierarchies with respect to the polynomial degree d, where each hierarchy is identified by a density f_0 and the type of the orthogonal polynomials used. A hierarchy reproduces and describes several temporal dynamics by estimating polynomial trends of different degrees. We refer to Dagum and Bianconcini (2006, 2008) for a theoretical study of properties of the RKHS-based filters by means of Fourier analysis.

4.6. Software

The Henderson filter is available in nonparametric seasonal adjustment software such as the X11 method developed by the U.S. Census Bureau (Shiskin et al., 1967), and its variants X-11-ARIMA (Dagum, 1980) and X-12-ARIMA (Findley et al., 1998). The LOESS filter is implemented in STL (Cleveland et al., 1990). Software implementing RKHS variants of Henderson and LOESS filters is available upon request. The Hodrick–Prescott filter exists in the most widely used statistical packages, e.g., Eviews, Stata, S-plus, R, Matlab, SAS.

5. SINGULAR SPECTRUM ANALYSIS

5.1. Preamble

In this section we consider the use of SSA for trend extraction. SSA is based on building a so-called trajectory matrix from a time series and on operating with the Singular Value Decomposition of this matrix. In some early references, this approach was referred to as the Karhunen–Loéve decomposition of a discrete time series. Based on the information provided by singular vectors, a matrix approximating the trajectory matrix is obtained, which is then converted into an additive component of the time series. Apart from the transformation of a time series to a matrix and an approximating matrix to a time series component, the algorithm of SSA coincides with the procedure of Principal Component Analysis.

SSA originated in the late 1970s and early 1980s, mainly in the area of dynamical systems, in particular as the result of Broomhead and King (1986) and Fraedrich (1986). The name SSA was introduced by Vautard and Ghil (1989), but this approach is also referred to as the *Caterpillar approach*; for surveys and references, see Golyandina et al. (2001) and Ghil et al. (2002). The similar ideas of SVD of the trajectory matrix have been used in other areas, e.g., for trend and cycles extraction by Basilevsky and Hum (1979) and for estimation of parameters of damped complex exponential signals (Kumaresan and Tufts, 1980). The present literature on SSA includes two monographs, several book chapters, and over a hundred articles.

SSA can be used in a wide range of issues: trend or periodical component extraction, denoising, forecasting, and change-point detection. At the present time, SSA is a well-known technique in the geosciences (Ghil and Vautard, 1991; Ghil et al., 2002), and it is starting to be applied in other areas, e.g., in biology, material science, nuclear science, and, recently, in economics (Hassani et al., 2009a,b; Thomakos, 2008).

5.2. Basic Algorithm and General Questions

The basic algorithm of SSA has two parts: *decomposition* of a time series and *reconstruction* of a desired additive component (e.g., a trend). For the decomposition, we choose a *window length* L $(1 < L < N)$ and construct a trajectory matrix $\mathbf{X} \in \mathbb{R}^{L\times K}$, $K = N - L + 1$, with stepwise portions of the time series x_n taken as columns:

$$(x_0, \ldots, x_{N-1}) \to \mathbf{X} = \begin{pmatrix} x_0 & x_1 & \ldots & x_{N-L} \\ x_1 & x_2 & \ldots & x_{N-L+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{L-1} & x_L & \ldots & x_{N-1} \end{pmatrix}$$

Then we perform the Singular Value Decomposition (SVD) of $\mathbf{X}$, where the j'th SVD component is described by an eigenvalue λ_j and a real-valued eigenvector U_j of $\mathbf{X}\mathbf{X}^{\mathrm{T}}$:

$$\mathbf{X} = \sum_{j=1}^{L} \sqrt{\lambda_j} U_j V_j^{\mathrm{T}}, \qquad V_j = \mathbf{X}^{\mathrm{T}} U_j \big/ \sqrt{\lambda_j}.$$

The SVD components are numbered in the decreasing order of their eigenvalues. The reconstruction stage combines (i) selection of a group $\mathcal{J}$ of several SVD components and (ii) reconstruction of a trend by hankelization (averaging along anti-diagonals) of the matrix formed from the selected part $\mathcal{J}$ of the SVD:

$$\sum_{j\in\mathcal{J}} \sqrt{\lambda_j} U_j V_j^{\mathrm{T}} \to \text{trend estimation}. \tag{7}$$

For a complete description of the algorithm, see (Golyandina et al., 2001, Chapter 1).

Trend extraction in SSA requires (i) choosing a window length L and (ii) selecting a group $\mathcal{J}$ of SVD components used for trend reconstruction. The former problem had only empirical solutions till Nekrutkin (1996), who showed how the quality of SSA decomposition depends on L. His *separability theory* provides instructions for choosing L according to the properties of assumed components of a time series, such as a trend, periodical components, and a noise (Golyandina et al., 2001, Chapter 6). Existing solutions for the second problem are considered further in Section 5.4.

5.3. Trend in SSA

SSA is essentially a nonparametric approach and does not need a priori specification of a model for a time series or for a trend, neither deterministic nor stochastic one. SSA can be used for trend extraction in two ways: (i) supposing separability between trend and residual and (ii) as a short smoothing nonparametric data-adaptive filter. First approach allows one to extract trend when a residual is supposed to have a complicated structure (containing not only noise, but also seasonal components or cycles). Second approach, similar to nonparametric filtering of Section 4, can separate a trend only from a simple residual (e.g., a white noise) and, hence, can be used only for trend-cycle extraction and requires a seasonally adjusted time series.

Extraction of a Trend Which is Separable from a Residual

The classes of trends and residual which can be successfully separated by SSA using the first approach are characterized as follows.

First, since we extract trend by selecting a subgroup of all L SVD components, the trend should generate only d $(d < L)$ of them. For infinite time series, the class of such trends coincides with the class of time series governed by finite difference equations (Golyandina et al., 2001). This class can be described explicitly as linear combinations of products of polynomials, exponentials, and sines. Naturally, an element of this class approximates well a smooth time series.

Second, a residual should be separable from a trend. The SSA-separability theory postulates the following: (i) every non-stochastic function can be asymptotically separated from any ergodic stochastic noise (Golyandina et al., 2001; Nekrutkin, 1996) as the time series length and the window length tends to infinity; and (ii) under the separability conditions a trend can be separated from a quasi-periodic component.

This approach provides flexible nonparametric framework for extraction of a trend from a time series of complex structure. However, this requires a careful choice of SVD components used for trend reconstruction. What is more, the quality of approximate separability increases as L grows that forces one to select L as large as possible, i.e., close to $N/2$ (and divisible by a seasonal period, if a time series includes a seasonal component). This increases the runtime of the algorithm.

SSA as a Short Nonparametric Smoothing Filter

Another approach to trend extraction using SSA is similar to nonparametric linear filtering described in Section 4. Given a small window length L and the SVD components used for reconstruction, SSA reconstruction at each time point can be seen as application of a linear filter. Its length is equal to L in the middle of time series and is smaller at the ends. The filter coefficients are determined from the data through the truncated SVD used in reconstruction 7.

This method provides a smoothing data-adaptive filter. Thomakos (2008) studied this filter for a unit-root random walk process, proved that this filter takes the form of particular moving average, and provided a frequency response function for it. Moreover, for a local level process, he compared his SSA-derived filter with Kalman fixed point smoothing for this process (McElroy, 2008) and showed that non-parametric SSA smoother is very competitive to the parametric Kalman filter, which is optimal for the data generating process and can be even better for a large signal-to-noise ratio.

Note that processing a time series, SSA considers all time series elements at once, for it considers the SVD of the trajectory matrix built from all parts of the time series. Therefore, SSA is not a local method in contrast to filter methods of Section 4. On the other hand, this makes SSA robust to outliers. Even for a short L, when SSA reconstruction is equivalent to application of a linear filter, the coefficients of this filter at each time point depend on the whole time series. However, in special cases SSA is a local method. Thomakos (2008) proved that SSA with the first SVD component used for trend reconstruction applied to a unit-root process is equivalent to a symmentric moving average filter with coefficients independent on the time series.

5.4. Methods of Trend Extraction in SSA

The naive idea of SVD components selection for trend extraction is to consider only eigenvalues but not eigenvectors and to take a few first SVD components. This simple approach works in many real-life cases on account of the optimal properties of SVD (Golyandina et al., 2001, Chapter 4). An eigenvalue represents the contribution of the corresponding SVD component into the trajectory matrix and into the original time series, respectively. Since trend characterizes the time series shape, its eigenvalues are larger than the others, which implies small order numbers for the trend SVD components. However, this selection procedure fails when the trend elements are small as compared with the residual (Golyandina et al., 2001, Section 1.6).

A more clever way of selection of trend SVD components is to choose those with smooth Empirical Orthogonal Functions (EOFs), where the

nth EOF is defined as the sequence of elements of the nth eigenvector. This approach is described in depth in (Golyandina et al., 2001), where the cases of polynomial and exponential trends are thoroughly examined. Using the concept of trajectory vector space which is spanned by the columns of the trajectory matrix and has the eigenvectors as an orthonormal basis, one can prove that the smoothness of a trend controls the smoothness of its EOFs on the assumption of separability of the trend and the residual.

The methods following this approach are described in (Alexandrov, 2009; Golyandina et al., 2001; Vautard et al., 1992). Vautard et al. (1992) exploited the Kendall correlation coefficient; for the properties of this correlation coefficient their method suits for extraction of monotonous trends only. Golyandina et al. (2001) proposed to select trend SVD components by visual examination of EOFs, which is a flexible approach but requires experience in application of SSA and understanding of the main principles of SSA representation of time series. Alexandrov (2009) presented a method based on the periodogram properties of EOFs. An original modification of SSA for producing smooth trends was proposed in (Solow and Patwardhan, 1996). Instead of calculating the eigensystem of $\mathbf{X}\mathbf{X}^{\mathrm{T}}$, the authors considered a special matrix built of the first differences of a time series.

5.5. Advantages and Disadvantages

SSA is a model-free approach that provides good results for short time series (Vautard et al., 1992), and allows one to extract trends from a wide class of time series. An essential disadvantage of SSA is the burden of computing of SVD. This cost can be reduced by using parallel computing or effective algorithms calculating only the first SVD components. For updating SVD in the case of receiving new time series in the future (trend revision), a computationally attractive algorithm of Gu and Eisenstat (1993) can be used. Moreover, Drmač and Veselić (2005) recently proposed a new method of SVD calculation which is as fast as QR-factorization and as stable as the conventional Jacobi method.

5.6. Software

The main software packages for trend extraction implementing SSA include SSA-MTM Toolkit (Vautard et al., 1992), CaterpillarSSA (Golyandina et al., 2001), AutoSSA (Alexandrov, 2009), and kSpectra Toolkit. SSA-MTM Toolkit implements a Kendall-correlation based method; CaterpillarSSA provides an interactive framework for time series processing and can be used for trend extraction and forecast; kSpectra Toolkit is a commercial version of SSA-MTM Toolkit. AutoSSA for Windows implements three parametric methods (in particular, the Low Frequencies (LF) and Kendall methods); AutoSSA for Matlab includes the LF method with adaptive selection of the parameters. Moreover, SSA is implemented in other scripts and packages; for details, see the website SSAwiki (http://www.math.uni-bremen.de/~theodore/ssawiki).

6. WAVELETS

6.1. Preamble

The term *wavelet* was introduced to the general scientific community in a pioneering article of Grossmann and Morlet (1984). However, a closer look at mathematical history (Meyer, 1993) reveals that several almost identical approaches—or at least similar concepts—have been around since the 1930s. A first wavelet construction can be found in several investigations of suitable model spaces for signals and functions (Littlewood–Payley theory). In addition, the Calderon's identity—or more recently the so-called pyramidal algorithms—share some features with wavelet methods. The term *wavelet analysis* is currently used for the somewhat larger field of multiscale analysis, with both its theoretical, mathematical foundations, and its resulting algorithms in signal and image processing.

A wavelet is a continuous square-integrable function on a finite support, which has a frequency content that is to be found, predominantly, within a specific band. A set of wavelet functions, ordered according to their temporal locations and by their frequency contents ("scales"), provides an orthogonal basis for continuous functions of time of an appropriately limited frequency content. From this perspective, wavelets analysis starts with associating an amplitude coefficient to each of the elements of the wavelets basis what can be efficiently done by means of *Finite Impulse Response* (FIR) filters.

Note that a subset of temporally-shifted wavelet functions has its own scale (frequency band). In this sense, expansion of a function in this basis splits the function into components of different scales (i.e., multiscale analysis). This property, as well as the existence of efficient algorithms of calculation of wavelet coefficients using FIR filters, and flexibility in choosing the mother wavelet generating the wavelet basis have led to several outstanding applications, including the design of efficient image and video compression standards (JPEG2000, MPEG) and advanced audio technology. In addition, the analysis of wavelet methods in a statistical framework has led to some powerful methods, e.g., for denoising signals and images (Donoho, 1995). Wavelet methods are now a generally accepted alternative to more classical statistical approaches or filtering techniques in a wide range of applications (see Vedam and

Venkatasubramanian, 1997; Partal and Kuecuek, 2006), and the citations therein.

Despite their success in engineering applications, wavelet methods have been only marginally used for trend extraction. However, recent works, see Bakshi and Stephanopoulos (1994), Vedam and Venkatasubramanian (1997), Bakhtazad et al. (2000), Tona et al. (2005), and Partal and Kuecuek (2006), in engineering literature and Pollock and Lo Cascio (2007) for an econometric viewpoint, revealed their potential in trend research.

6.2. Basic Algorithms and General Considerations

The basic wavelet algorithm computes a decomposition (wavelet transform) of a time series into wavelet coefficients with a pair of FIR filters in several steps. Let us introduce some additional notation and call the original time series $c^0 = \{c_k^0 \mid k = 0, 1, \ldots, N-1\}$. In a first step, the data c^0 is convolved with a low pass and a band pass filters $\{h_\ell\}$ and $\{g_\ell\}$, respectively. The results of filtering are then subsampled by factor of 2:

$$c_k^1 = \sum_\ell h_{2k-\ell} c_\ell^0, \qquad d_k^1 = \sum_\ell g_{2k-\ell} c_\ell^0.$$

Given FIR filters $\{h_\ell\}$ and $\{g_\ell\}$ corresponding to a wavelet basis, the sequence of coefficients d^1 represents the wavelet coefficients of the finest scale (usually corresponding to a noise). The low pass filter generates a smoothed version c^1 of the original time series. At this step, d^1 is stored and c^1 is passed to the second step. At the second step, all computations of the first step are repeated on a low pass filtered version c^1. The outcome of the subsampled band pass filtering in the second step (d^2) is again stored, it contains the details of the original time series on scale two. The application of the subsampled low pass filter gives an even smoother version of the original time series.

Repeating this process for a fixed number of s steps produces a family of sequences, which represent details on different scales $\{d^j, j = 1, \ldots, s\}$ as well as a very smooth final version c^s:

$$c_k^{j+1} = \sum_\ell h_{2k-\ell} c_\ell^j, \qquad d_k^{j+1} = \sum_\ell g_{2k-\ell} c_\ell^j, \qquad j = 0, 1, \ldots, s-1.$$

Note that due to the subsampling the overall number of wavelet coefficients, stored in sequences $\{d^j, j = 1, \ldots, s\}$ and c^s, is equivalent to the length of the time series c^0.

The success of wavelet methods relies on the existence of a dual pair of low and band pass filters which reconstruct the original time series c^0 given $\{d^j, j = 1, \ldots, s\}$, and c^s:

$$c_k^j = \sum_\ell h_{k-2\ell} c_\ell^{j+1} + \sum_\ell g_{k-2\ell} d_\ell^{j+1}, \qquad j = s-1, s-2, \ldots, 0.$$

The filters need to satisfy some stability criteria in order to control reconstruction errors. The choice of an appropriate filter bank for decomposition and reconstruction is crucial for the success of wavelet methods. There exists an extensive library of wavelet filters that are appropriate for many kinds of applications.

In principle, a wavelet decomposition algorithm is fully defined by the choice of the wavelet filters and the number of decomposition steps s. However, we want to emphasize that before using a wavelet algorithm, one should answer the following two basic questions: (i) Why should we use a wavelet method? and (ii) How do we want to analyze the wavelet decomposition?

The first question can be answered positively whenever the time series under consideration has a multiscale (as opposed to a multifrequency) structure. This in particular includes the analysis of non-stationary effects, e.g., defects in signals for monitoring technical processes or applications of change-of-trend detection.

The second question is extremely important. The wavelet transform only generates an alternative representations of the time series: no information is lost, no information is added by the wavelet transform. For certain applications, however, we might expect that any sought-after information can be more easily detected in the transformed data. Hence, at the beginning we need to consider how we want to extract this information after having computed the wavelet transform.

6.3. Trend Extraction with Wavelet Methods

Following the general considerations described in the previous subsection, we first need to determine why wavelet methods should be useful for trend analysis. In this section we follow the model (2) with trend-cycle component t_n, seasonal component d_n, and noise-component s_n. This model is well-suited for a wavelet decomposition, which results in an additive multiscale decomposition $\{c^s, d^j, j = 1, \ldots, s\}$ of the time series x_n. An increasing "scale of detail" is assigned to every component of the decomposition, i.e., we interpret the different components as being the sum of all details in the time series which live on a prescribed scale (resolution or size of detail).

Therefore, a typical wavelet decomposition concentrates the noise component on the first fine scales, the trend component on the coarse scale (c^s) and other components of the signal (cyclic or seasonal) produce coefficients d^j on the intermediate scales. Note that in the conventional wavelet analysis a seasonal or monthly component cannot be represented by a set of wavelet coefficients of a single scale as shown by Pollock and Lo Cascio (2007) who also proposed non-dyadic wavelet analysis to solve this issue.

Hence, the basic trend extraction procedure with wavelet methods proceeds by (i) choosing an appropriate wavelet filter bank, (ii) computing a wavelet decomposition up to scale s, (iii) deleting all fine scales (scales of noise and seasonal components), and (iv) reconstructing the remaining additive component.

Some of the prominent applications for trend extraction by wavelet methods include process monitoring of technical processes (Bakhtazad et al., 2000; Vedam and Venkatasubramanian, 1997) and analysis of environmental data (Partal and Kuecuek, 2006; Tona et al., 2005). This list is neither complete nor representative, but rather serves to demonstrate some of basic examples presented in the vast literature on wavelet trend analysis.

This typically results in a good "visualization" of the underlying non-stationary trend. In this sense, wavelet methods use a semiparametric trend model: the choice of the wavelet filters determines the trend model, since it determines whether we capture piecewise constant, linear, polynomial or exponential trends (see Bakshi and Stephanopoulos, 1994) or the general references on polynomial reproduction by wavelet basis in Louis et al. (1997) and Mallat (2001). On the other hand, the choice of the wavelet is only important for the intermediate computations. After reconstruction, we obtain a trend model in the physical space given by the measurement data. The trend can be subsequently analyzed without any underlying model.

We want to emphasize that applying a wavelet method usually constitutes just one step in a more complex scenario for trend analysis. Typical tasks, e.g., change point detection, require to analyze the extracted trend and to give precise estimates for the time instances of change-of-trend. There exist refined wavelet methods based on shrinkage operations for this kind of analysis. Change-of-trend features exist on all scales of resolution; hence these methods use the full wavelet decomposition and rely on adaptive thresholding procedures on all scales (Mallat, 2001; Partal and Kuecuek, 2006; Vedam and Venkatasubramanian, 1997).

A different scenario for the application might require the determination of a physical model for the underlying process. In this case, the trend extraction and the determination of different time intervals with a stationary behavior are only a first step. Hence, the characterization of the model on each time interval with a stable trend is then left to other methods of signal analysis (e.g., dynamical systems or the methods following MBA).

To conclude, the potential of wavelet methods for trend extraction is based on its non-stationary, quasi-local properties. Wavelet methods are well-suited for the determination of change-of-trend points, as well as the decomposition of the time axis in different time intervals with stable trend behavior.

6.4. Advantages and Disadvantages

The application of a wavelet algorithm starts by choosing an appropriate wavelet basis or wavelet filter bank. This offers some flexibility for optimization, since there exist highly specialized wavelet filters for a large variety of complex situations. Its main feature is a non-stationary multiscale decomposition, which is particularly suited for analyzing localized effects. This flexibility, which allows one to finely tune the wavelet algorithms to different specific tasks, is also one of its major disadvantages. However, some experience with wavelet methods is required in order to fully exploit its power.

In addition, the treatment of boundary effects can be crucial, especially when the data are trending. However, note that this problem is not specific to the wavelet analysis but occurs because wavelet algorithms are based on a repeated application of linear filters, which requires an adjustment at the end of the given data series. For example, the Fourier analysis is subject to this problem; see as well Section 3 on discussion of this problem in the context of MBA. The predominant approach to deal with boundary effects in the wavelet analysis is to add a sufficient amount of data points (either zeros or using the periodic continuation). A more advanced approach is to adapt the filter coefficients. The third approach is to extract a simple (polynomial or exponential) trend and to analyze the residual sequence.

Also, the visualization of wavelet transforms has not yet been fully standardized. Again, some experience is required in order to "understand" the results of the wavelet transform. Typically, the best way to analyze a wavelet transform is to reconstruct the manipulated wavelet decomposition and to display the result in the physical domain of the original signal.

The major advantage is the efficiency of the fast wavelet decomposition which is an $\mathcal{O}(N)$-algorithm. Efficient implementations are by now included in any software toolbox for signal or image analysis, e.g., MATLAB or S+.

7. SUMMARY TABLE AND A REAL-LIFE EXAMPLE

In this section we present Table 1 which summarizes the approaches considered in this article and apply them to a real-life example. The aim of the example is not to compare the resulting trends but to show a usual sequence of steps which a user must perform while applying one of the trend extraction methods.

Let us consider a time series of the electric power use by industry in the United States for the period from 1972/1 to 2005/10 provided by the Federal Reserve Board (FRB). The data is monthly of length 406, and is publicly available at: http://www.federalreserve.gov/releases/g17/ipdisk/kwh_nsa.txt. This time series was selected because: (i) it contains a clear and complex trend, (ii) we can test trend extraction in the presence of a sizeable seasonal component, (iii) the length of several hundred points is usual for many applications, and (iv) the noise is significant enough to demonstrate the smoothing properties of the various methods.

7.1. Application of Trend Methods

The time series has an evident seasonal component. MBA, SSA, and wavelets can extract trend from such data allowing a residual to be of complex structure (e.g., to contain a seasonal component), while nonparametric predictors (Henderson, LOESS, HP) require seasonally adjusted data. Fortunately, FRB provides seasonally adjusted time series, and we have applied the nonparametric methods to the seasonally adjusted variant of the same time series. The resulting trends are shown in Fig. 1, shifted for better visualization, each with the initial time series in background.

MBA. The regARIMA modeling of X-12-ARIMA software was used to identify regression parameters and significance, and determine the best model using AICc (Akaike information criterion with a second order correction for small sample sizes). This was a (1 1 0)(0 1 1) SARIMA model, and the canonical decomposition into trend, seasonal, and irregular exists. The minimum MSE trend extraction filter for finite sample (McElroy, 2008) was determined and applied.

SSA. We exploited a method of trend extraction of Alexandrov (2009) implemented in AutoSSA software for Matlab, available at: http://www.pdmi.ras.ru/~theo/autossa. First, we performed seasonal adjustment and then extracted the trend. For seasonal adjustment, we used the Fourier method for extraction of periodical components also implemented in AutoSSA. The window length is set to $L = 192$ (close to $N/2$ and divisible by a seasonal period 12). The used value of the low-frequency boundary is 0.07 (slightly smaller the seasonal frequency 1/12).

TABLE 1 Summary characteristics of the considered approaches to trend extraction

	Trend model	Residual model	Trend-residual regulations	Pre-specification	Software	Pros	Cons
MBA	Stochastic (ARIMA, SARIMA, structural model), allowing for deterministic polynomial portion	Typically a combination of cycle, irregular, and seasonal components given by stochastic models	Uncorrelation (orthogonal approach) or "full" correlation (BN approach) between differenced trend and residual	(1) Decomposition versus structural versus BN; (2) component models (differencing operators, ARMA models, etc.)	MicroCAPTAIN, STAMP, TRAMO-SEATS, X-12-ARIMA	The most developed approach to trend extraction, many theoretical results and methods	Parametric; only linear models are typically used; poorer results when models are misspecified
Nonparametric	Henderson, LOESS: trend is globally smooth and locally approximated by a polynomial; Hodrick-Prescott: no model (but see Section 4.4)	A stationary and invertible ARMA process, usually $NID(0, \sigma^2)$	Trend and residual are supposed to be uncorrelated in a specially-defined sense (see Eq. (3) and explanations afterwards)	Henderson and LOESS: filter span width; LOESS: polynomial degree, weight function; HP: only penalty coefficient	Henderson, LOESS: STL, X-11-ARIMA, X-12-ARIMA; HP: many toolboxes	Fast, simple, a few prespecifications is required	A residual of a complex structure is not allowed; only seasonally adjusted data; revisions when asymmetric filters are applied to the most recent observations
SSA	Large window length L: deterministic (finite rank time series); short L: no model	Large L: typically a combination of cycle and seasonal components with varying amplitudes plus irregular component; short L: irregular component	Large L: trend is to be separable (at least approximately) from the residual; short L: no regulations	Original SSA: L, SVD components; AutoSSA: L, trend frequency interval; SSA-MTM/kSpectra Toolkit: L, Kendall significance level	AutoSSA, CaterpillarSSA, kSpectra Toolkit, SSA-MTM Toolbox	A few prespecificaitons, can separate a trend from a complex residual, good for time series with a large noise	Few theoretical studies of trend estimators, computational complexity of SVD calculation; small L: seasonally adjusted data
Wavelets	Semi-parametric, specified by the wavelet	Very general, could include a combination of cycles and irregular component	Trend and residual are to have different scale or, that is the same for a special choice of the wavelet, different smoothness	Wavelet basis (or wavelet filters), levels used for trend reconstruction	Wavelet trans-formation and reconstruction: various software packages; no special wavelet- based trend extraction software	Efficient algorithms, many available wavelet bases are available, good smoothing properties	Subjective choice of levels used for trend reconstruction; boundary effects; no ready-made method for trend extraction

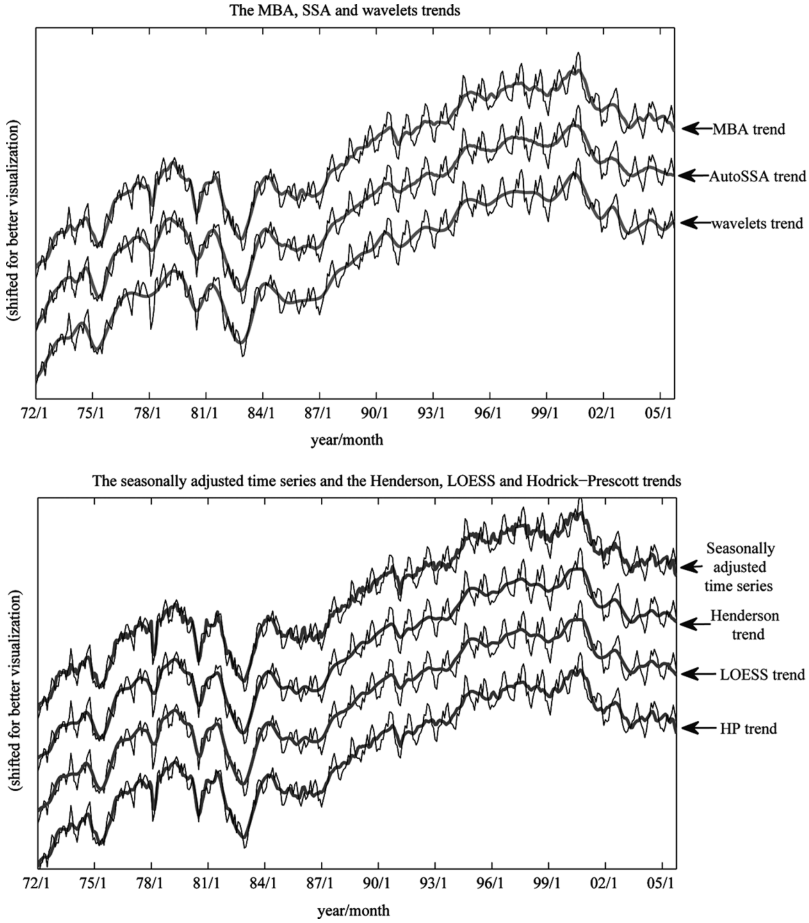

FIGURE 1 Top panel: The resulting MBA, SSA, and wavelet trends, shifted for better visualization, each with the initial time series in background; Bottom panel: The resulting Henderson, LOESS, and Hodrick–Prescott trends, as well as the seasonally adjusted variant of the time series, shifted for better visualization, each with the initial time series in background.

Wavelets. The Coifman wavelet of order 4 (*coif4*) was selected for its symmetry and good smoothing properties. After wavelet transformation, a trend was reconstructed by all wavelet coefficients excepting detail coefficients on levels 1 and 2. As a multiscale approach, wavelet transformation can extract trends of different resolution. The extracted trend seems to contain some insignificant portions of the seasonal component opposite to a trend reconstructed without details on levels 1, 2, and 3. Nevertheless, we selected the former one because it better represents the sought-for trend-cycle.

Nonparametric. Note that for nonparametric filtering we used the seasonally adjusted time series. The length of Henderson and LOESS filters was selected according to the signal-to-noise ratio equal to 1.14 (provided by X-11-ARIMA software). Hence, a 13-term filter is appropriate for the estimation of the trend-cycle. The trend estimates were obtained based on the following RKHS filters: (i) 13-term 3rd order Henderson kernel within the biweight hierarchy and (ii) 13-term 3rd order LOESS kernel within the tricube hierarchy. An HP trend was produced using the *pspline* R-package, with the smoothing parameter selected by means of generalized cross-validation.

7.2. Comparative Analysis of the Trends

Note that we do not aim at selecting the best method, but rather illustrate application of the methods. However, it is worth mentioning that the LOESS trend is very similar to the Henderson one, although the former has slightly larger values throughout the whole timespan except for boundaries (approximately by 0.3, cannot be seen in Fig. 1 since the trends are shifted). Moreover, Fig. 1 shows that the wavelet trend has slight upward bias at the end of the series, that probably demonstrates the boundary effect discussed in Section 6.4. Subjectively, the HP trend with the parameter selected using generalized cross-validation is not smooth enough. Note that in this example including more scales in wavelets leads to inclusion of a portion of the seasonal component into the trend; selecting a larger low-frequencies-boundary in AutoSSA reduces the smoothness of the trend considerably. In the early part of the series the MBA trend is arguably too oscillatory—especially as compared to the AutoSSA and wavelets methods—which is probably due to cyclical effects that are visually apparent in the data; a smoother trend could presumably be produced, if a user so desired, by explicitly modeling the cyclical effects and placing these in the residual component.

ACKNOWLEDGMENTS

The authors thank David F. Findley for comments on the manuscript, Nina Golyandina for discussion on SSA, and Simone Giannerini for the proofreading.

REFERENCES

Alexandrov, T. (2009). A method of trend extraction using singular spectrum analysis. *RevStat* 7(1):1–22.

Anderson, B., Moore, J. (1979). *Optimal Filtering*. Englewood Cliffs: Prentice-Hall.

Aronszajn, N. (1950). Theory of reproducing kernels. *Transaction of the AMS* 68:337–404.

Bakhtazad, A., Palazoglu, A., Romagnoli, J. (2000). Process trend analysis using wavelet-based de-noising. *Control Eng. Pract.* 8(6):657–663.

Bakshi, B., Stephanopoulos, G. (1994). Representation of process trendsiii: Multiscale extraction of trends rom process data. *Comput. Chem. Eng.* 18(4):267–302.

Basilevsky, A., Hum, D. P. J. (1979). Karhunen-Loeve analysis of historical time series with an application to plantation births in Jamaica. *J. Am. Stat. Assoc.* 74:284–290.

Bell, W. (1984). Signal extraction for nonstationary time series. *Ann. Stat.* 12:646–664.

Bell, W. (2004). On RegComponent time series models and their applications. In: Harvey, A. C., Koopman, S. J., Shephard, N., eds. *State Space and Unobserved Component Models: Theory and Applications*. Cambridge, UK: Cambridge University Press, pp. 248–283.

Bell, W., Hillmer, S. (1984). Issues involved with the seasonal adjustment of economic time series. *J. Bus. Econ. Stat.* 2:291–320.

Berlinet, A. (1993). Hierarchies of higher order kernels. *Probab. Theory Rel.* 94:489–504.

Berryman, D., Bobee, B., Cluis, D., Haemmerli, J. (1988). Nonparametric tests for trend detection in water quality time series. *Water Resour. Bull.* 24(3):545–556.

Beveridge, S., Nelson, C. (1981). A new approach to decomposition of economic time series into permanent and transitory components with particular attention ro measurement of the business cycle. *J. Monetary Econ.* 7:151–174.

Brockwell, P. J., Davis, R. A. (2003). *Introduction to Time Series and Forecasting*. New York: Springer.

Broomhead, D. S., King, G. P. (1986). Extracting qualitative dynamics from experimental data. *Physica D* 20:217–236.

Burman, J. (1980). Seasonal adjustment by signal extraction. *J. R. Stat. Soc. Ser. A–G* 143:321–337.

Chatfield, C. (1996). *The Analysis of Time Series: An Introduction.* Boca Raton, FL: Chapman & Hall/CRC.

Cleveland, R., Cleveland, W., McRae, J., Terpenning, I. (1990). STL: A seasonal trend decomposition procedure based on LOESS. *Journal of Official Statistics* 6(1):3–33.

Cleveland, W. (1979). Robust locally regression and smoothing scatterplots. *J. Am. Stat. Assoc.* 74:829–836.

Cleveland, W., Tiao, G. (1976). Decomposition of seasonal time series: A model for the Census X-11 program. *J. Am. Stat. Assoc.* 71:581–587.

Dagum, E. (1980). *The X11ARIMA Seasonal Adjustment Method*. Ottawa: Statistics Canada Publication, Catalogue No. 12-564E.

Dagum, E., Bianconcini, S. (2006). Local polynomial trend-cycle predictors in Reproducing Kernel Hilbert Spaces for current economic analysis. Proceedings of the International Congress of Applied Economics (ASEPELT 2006, La Laguna, Spain). *Anales de Economia Aplicada*, pp. 1–22.

Dagum, E., Bianconcini, S. (2008). The Henderson smoother in reproducing Kernel Hilbert space. *Journal of Business and Economic Statistics* 26(4):536–545.

Donoho, D. (1995). Denoising by soft-thresholding. *IEEE Trans. Inf. Theory* 41(3):613–627.

Drmač, Z., Veselić, K. (2005). *New Fast and Accurate Jacobi svd Algorithm: I, II.* Technical Report LAPACK Working Note 169, Dep. of Mathematics, University of Zagreb, Croatia.

Durbin, J., Koopman, S. J. (2001). *Time Series Analysis by State Space Methods*. Volume 24 of *Oxford Statistical Science Series*. Oxford: Oxford University Press.

Esterby, S. R. (1996). Review of methods for the detection and estimation of trends with emphasis on water quality applications. *Hydrol. Process.* 10:127–149.

Findley, D. F., Monsell, B. C., Bell, W. R., Otto, M. C., Chen, B. C. (1998). New capabilities and methods of the X-12-ARIMA seasonal adjustment program. *J. Bus. Econ. Stat.* 16(2):127–177.

Fraedrich, K. (1986). Estimating dimensions of weather and climate attractors. *J. Atmos. Sci.* 43:419–432.

Froeb, L., Koyak, R. (1994). Measuring and comparing smoothness in time series: The production smoothing hypothesis. *J. Econometrics* 64(1–2):97–122.

Gersch, W., Kitagawa, G. (1983). The prediction of time series with trends and seasonalities. *J. Bus. Econ. Stat.* 1(3):253–264.

Ghil, M., Vautard, R. (1991). Interdecadal oscillations and the warming trend in global temperature time series. *Nature* 350:324–327.

Ghil, M., Allen, R. M., Dettinger, M. D., Ide, K., Kondrashov, D., Mann, M. E., Robertson, A., Saunders, A., Tian, Y., Varadi, F. P. Y. (2002). Advanced spectral methods for climatic time series. *Rev. Geophys.* 40(1):1–41.

Golyandina, N. E., Nekrutkin, V. V., Zhigljavsky, A. A. (2001). *Analysis of Time Series Structure: SSA and Related Techniques*. Boca Raton, FL: Chapman & Hall/CRC.

Grossmann, A., Morlet, J. (1984). Decomposition of Hardy functions into square integrable wavelets of constant shape. *SIAM J. of Math. Anal.* 15(4):723–736.

Gu, M., Eisenstat, S. C. (1993). *A Stable and Fast Algorithm for Updating the Singular Value Decomposition.* Technical Report YALEU/DCS/RR-966, Yale University.

Hamed, K., Rao, A. (1998). A modified Mann-Kendall trend test for autocorrelated data. *J. Hydrol.* 204(1):182–196.

Harvey, A. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.

Harvey, A., Jaeger, A. (1993). Detrending, stylized facts and the business cycle. *Journal of Applied Econometrics* 8:231–247.

Hassani, H., Heravi, S., Zhigljavsky, A. (2009a). Forecasting European industrial production with singular spectrum analysis. *International Journal of Forecasting* 25(1):103–118.

Hassani, H., Soofi, A. S., Zhigljavsky, A. A. (2009b). Predicting daily exchange rate with singular spectrum analysis. *Nonlinear Analysis: Real World Applications* 11(3):2023–2034.

Henderson, R. (1916). Note on graduation by adjusted average. *Trans. Actuar. Soc. Amer.* 17:43–48.

Hillmer, S., Tiao, G. (1982). An ARIMA-model-based approach to seasonal adjustment. *J. Am. Stat. Assoc.* 77(377):63–70.

Hirsch, R. M., Slack, J. R. (1984). Nonparametric trend test for seasonal data with serial dependence. *Water Resources Research* 20(6):727–732.

Hodrick, R., Prescott, E. (1997). Postwar U.S. business cycles: An empirical investigation. *Journal of Money, Credit and Banking* 29(1):1–16.

Kaiser, R., Maravall, A. (2001). *Measuring Business Cycles in Economic Time Series.* New York: Springer-Verlag.

Kalman, R. (1960). A new approach to linear filtering and prediction problems. *J. Basic Eng-t ASME D* 82:35–45.

Kenny, P., Durbin, J. (1982). Local trend estimation and seasonal adjustment of economic and social time series. *Journal of the Royal Statistical Sociaety A* 145:1–41.

King, R., Rebelo, S. (1993). Low frequency filtering and real business cycles. *Journal of Economic Dynamics and Control* 17:207–233.

Kitagawa, G., Gersch, W. (1996). *Smoothness Priors Analysis of Time Series.* New York: Springer.

Koopman, S. (1997). Exact initial Kalman filtering and smoothing for nonstationary time series models. *J. Am. Stat. Assoc.* 92:1630–1638.

Koopman, S., Harvey, A., Doornik, J., Shepherd, N. (2000). *Stamp 6.0: Structural Time Series Analyser, Modeller, and Predictor.* London: Timberlake Consultants.

Koopman, S., Shepherd, N., Doornik, J. (1999). Statistical algorithms for models in state space using SsfPack 2.2. *Economet. J.* 2:113–166.

Koopmans, L. (1974). *The Spectral Analysis of Time Series.* New York: Academic Press.

Kumaresan, R., Tufts, D. W. (1980). Data-adaptive principal component signal processing. In Proc. of IEEE Conference On Decision and Control. Albuquerque, pp. 949–954.

Louis, A., Maass, P., Rieder, A. (1997). *Wavelets: Theory and Applications.* West Sussex, UK: Wiley.

Mallat, S. (2001). *A Wavelet Tour of Signal Processing*. 2nd ed. San Diego, CA: Academic Press.

Maravall, A., Caporello, G. (2004). *Working Article 2004. Program TSW: Revised Reference Manual.* Technical Report, Research Dep., Bank of Spain. http://www.bde.es.

McElroy, T. (2008). Matrix formulas for nonstationary signal extraction. *Econometric Theory* 24:1–22.

Meyer, Y. (1993). *Wavelets: Algorithms and Applications.* Philadelphia: SIAM.

Musgrave, J. (1964). *A Set of End Weights to End All End Weights.* Technical Report, US Census Bureau.

Nekrutkin, V. (1996). Theoretical properties of the "Caterpillar" method of time series analysis. In Proc. 8th IEEE Signal Processing Workshop on Statistical Signal and Array Processing. IEEE Computer Society, pp. 395–397.

Partal, T., Kuecuek, M. (2006). Long-term trend analysis using discrete wavelet components of annual precipitation measurements in the marmara region. *Phys. Chem. of the Earth* 31(18):1189–1200.

Peña, D., Tiao, G., Tsay, R. (2001). *A Course in Time Series Analysis.* New York: Wiley.

Pierce, D. (1980). Data revisions with moving average seasonal adjustment procedures. *J. Econometrics* 14:95–114.

Pollock, D. (2007). Weiner-kolmogorov filtering, frequency-selective filtering and polynomial regression. *Econometric Theory* 23:71–83.

Pollock, D. S. G. (2008). Investigating econometric trends and cycles. In: Mills, T. C., Patterson, K., eds. *Palgrave Handbook of Econometrics: Vol. 2 Applied Econometrics*. Basingstroke, UK: Palgrave Macmillan, pp. 243–307.

Pollock, D. S. G., Lo Cascio, I. (2007). Non-dyadic wavelet analysis. In: Kontogoghiorges, E., Gatu, C., eds. *Optimisation, Econometric and Financial Analysis*. Berlin, Germany: Springer-Verlag, pp. 167–203.

Proietti, T. (2006). Trend-cycle decompositions with correlated components. *Economet. Rev.* 25:61–84.

Rauch, H. (1963). Solution to the linear smoothing problem. *IEEE Transactions on Automatic Control* 8:371–372.

Ravn, M., Uhlig, H. (1997). *On Adjusting the HP-filter for the Frequency of Observations*. Technical report, Tilburg University.

Reinsch, C. (1967). Smoothing by spline functions. *Numerische Mathematik* 10:177–183.

Shiskin, J., Young, A., Musgrave, J. (1967). The X11 variant of the Census method II seasonal adjustment program. *Technical Paper 15, U.S. Department of Commerce, Bureau of the Census, Washington, D.C.*

Slepian, D. (1976). On bandwidth. *Proc. of the IEEE* 64(3):292–300.

Solow, A. R., Patwardhan, A. (1996). Extracting a smooth trend from a time series: A modification of singular spectrum analysis. *J. Climate* 9:2163–2166.

Thomakos, D. (2008). *Optimal Linear Filtering, Smoothing and Trend Extraction for Processes with Unit Roots and Cointegration*. Working Paper 0024, University of Peloponnese, Department of Economics. Available at http://ideas.repec.org/p/uop/wpaper/0024.html.

Thomson, P., Ozaki, T. (2002). Transformation and trend-seasonal decomposition. In Proc. of the 3rd Int. Symp. on Frontiers of Time Series Modeling. Tokyo, Japan.

Tona, R., Benqlilou, C., Espuna, A., Puigjaner, L. (2005). Dynamic data reconciliation based on wavelet trend analysis. *Ind. Eng. Chem. Res.* 44(12):4323–4335.

Vautard, M., Ghil, M. (1989). Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series. *Physica D* 35:395–424.

Vautard, R., Yiou, P., Ghil, M. (1992). Singular-spectrum analysis: A toolkit for short, noisy chaotic signals. *Physica D* 58:95–126.

Vedam, H., Venkatasubramanian, V. (1997). A wavelet theory-based adaptive trend analysis system for process monitoring and diagnosis. In *Proc. Amer. Cont. Conf.* Albuquerque, NM, pp. 309–313.

Wahba, G. (1990). *Spline Models for Observational Data*. Philadelphia: SIAM.

Whittle, P. (1963). *Prediction and Regulation*. London: English Universities Press.

Wiener, N. (1949). *The Extrapolation, Interpolation, and Smoothing of Stationary Time Series With Engineering Applications*. New York: Wiley.

Wildi, M. (2005). *Signal Extraction*. Berlin: Springer-Verlag.

Wold, H. (1938). *A Study in the Analysis of Stationary Time Series*. Stockholm, Sweden: Almquist & Wiksells.

Young, P. (1984). *Recursive Estimation and Time Series Analysis*. New York: Springer-Verlag.

Young, P., Benner, S. (1991). *MicroCAPTAIN Handbook: Version 2.0*. Center for Research on Environmental Systems and Statistics, Lancaster University.

Young, P., Pedregal, D. (1999). Recursive and en-block approaches to signal extraction. *J. Appl. Stat.* 26:103–128.

Young, P., Pedregal, D., Tych, W. (1999). Dynamic harmonic regression. *J. Forecasting* 18:369–394.

Yue, S., Pilon, P. (2004). A comparison of the power of the *t*-test, Mann–Kendall and bootstrap tests for trend detection. *Hydrol. Sci. J.* 49(1):21–37.