Computational Statistics & Data Analysis 42 (2003) 219–241
www.elsevier.com/locate/csda

Jump process for the trend estimation of time series

Shan Zhao[a], G.W. Wei[a,b,*]

[a] *Department of Computational Science, National University of Singapore, Singapore 117543, Singapore*
[b] *Department of Mathematics, Michigan State University, East Lansing, MI 48824, USA*

Received 1 January 2001; received in revised form 1 March 2002; accepted 1 March 2002

Abstract

A jump process approach is proposed for the trend estimation of time series. The proposed jump process estimator can locally minimize two important features of a trend, the smoothness and fidelity, and explicitly balance the fundamental tradeoff between them. A weighted average form of the jump process estimator is derived. The connection of the proposed approach to the Hanning filter, Gaussian kernel regression, the heat equation and the Wiener process is discussed. It is found that the weight function of the jump process approaches the Gaussian kernel, as the smoothing parameter increases. The proposed method is validated through numerical applications to both real data analysis and simulation study, and a comparison with the Henderson filter.

Keywords: Jump process; Time series; Trend estimation; Nonparametric regression; The smoothness-fidelity tradeoff; Weighted average form; Gaussian kernel

1. Introduction

In the additive model of the time series, it is assumed that the economic time series is made up of three components, the trend, the cyclic (seasonal) component, and the irregular component (random noise),

$$y_t = T_t + S_t + \varepsilon_t, \quad t = 1, 2, \ldots, N, \tag{1}$$

where random variable y_t is the observation at a discrete time t, T_t and S_t are trend and seasonal component, respectively, and N is the length of the time series. Here, ε_t

* Corresponding author.
E-mail address: cscweigw@nus.edu.sg (G.W. Wei).

0167-9473/02/$ - see front matter
PII: S0167-9473(02)00125-1

is random noise and is usually assumed to be stationary and of zero mean. From the viewpoint of mathematical modeling, trend is not a well-defined concept, see Kenny and Durbin (1982), Harvey (1989), Chatfield (1996), Mosheiov and Raveh (1997), and Franses (1998). In general, trend may be considered as "long-term smooth change in the mean level".

To carry out in-depth statistical study of time series, it is often necessary to convert a nonstationary series into a stationary one before the statistical model is treated. In other words, time series will be decomposed into individual trend, seasonal, and irregular components. Conventionally, there are two different ways to decompose a time series which does not consist of seasonal component (or is seasonally adjusted): differencing and detrending. The differencing will remove trend, while the detrending (trend estimation) will present an estimate to trend. Thus, the differencing and detrending are essentially high-pass filter and low-pass filter, respectively, from the viewpoint of digital signal processing (DSP). It is noted that a debate has been arisen about the appropriate selection between these two alternative ways for economic and financial time series, and the issue remains unresolved, see Nelson and Plosser (1982), Sims (1988), Sims and Uhlig (1991), Campbell and Perron (1991), Cochrane (1991), and the special issue of Vol. 6 (1991) of the Journal of Applied Econometrics.

Since the nonstationary trend component sometimes can be of more interest than the noise, the trend estimation techniques have attracted a lot of research interests in the literature, see for example, Kenny and Durbin (1982), Goodall (1990), Ball and Wood (1996), Mills and Crafts (1996), Mosheiov and Raveh (1997), Canova (1998), Blanchi et al. (1999), Wen and Zeng (1999), Ferreira et al. (2000), and Pollock (2000). In fact, the trend estimation is potentially useful for data interpretation, long-term forecasting and for the study of real business cycle. In some cases, the detection of trend is even a crucial task, see for example Visser and Molenaar (1995). Therefore, it is practically useful to accurately estimate the trend component of time series.

Generally speaking, methods of trend estimation fall into two major categories: parametric and nonparametric. In the parametric approach, a deterministic trend is commonly expressed by a particular smoothing function or model, such as a polynomial, the Gompertz curve or the logistic curve (Meade and Islam, 1995), or the structural time series model (Harvey, 1989). However, the use of an inappropriate parametric model may cause misleading information and even incorrect inference about the trend curve. Therefore, alternative nonparametric trend estimation methods are widely used. In particular, nonparametric approaches offer considerable flexibility in the selection of fitting curves and may yield satisfactory estimates. In conventional time series analysis, some of the most widely used nonparametric approaches include moving average filters and exponential smoothing filter, which are linear low-pass filters in the sense of the DSP. A variety of moving average filters are proposed, such as Spencer filter (Kendall et al., 1983), Henderson filter (Kenny and Durbin, 1982), and GLAS filter (Blanchi et al., 1999), etc. The asymmetric exponential smoothing filter (Kenny and Durbin, 1982) has a distinguished advantage in the treatment of boundary effect, hence is often preferred for the purpose of forecasting.

Apart from these linear filters, various nonparametric regression estimators existing in the literature can be easily adopted for the purpose of trend estimation. This is because that the counterpart of trend component T_t in the content of nonparametric regression is just the regression function $T(x_t)$,

$$y_t = T(x_t) + \varepsilon_t, \quad t = 1, \ldots, N, \tag{2}$$

where x and y are explanatory and response variable, respectively. Many powerful nonparametric regression estimators have been proposed and applied in the literature, such as kernel smoothing (Müller, 1988), LOESS (Cleveland, 1979), locally weighted polynomial regression (Fan and Gijbels, 1996), smoothing spline (Eubank, 1999), and regression spline (Doksum and Koo, 2000). Furthermore, there are also many interesting studies suggested in the literature in order to enhance the performance of nonparametric regression estimators, such as improving the accuracy (e.g., Borra and Ciaccio, 2002), dealing with special data (e.g., Keilegom et al., 2001), and so on. In fact, the nonparametric regression has been successfully used in trend estimation, see for example Härdle and Tuan (1986), Goodall (1990), Hart (1991, 1994), Høst (1999), and Ferreira et al. (2000). Other nonparametric trend estimation methods that were studied in the literature include: Hodrick Prescott (HP) filter (Hodrick and Prescott, 1997), median filter (Wen and Zeng, 1999), wavelet shrinkage (Donoho et al., 1995), and linear programming (Mosheiov and Raveh, 1997). Usually, a successful nonparametric method has one or more underlying smoothing parameters which can be adjusted to balance the fundamental tradeoffs of the estimates, i.e., the smoothness-fidelity tradeoff and the variance-bias tradeoff.

What is more relevant to the present work is a class of nonparametric trend estimation methods that attempt to globally quantify the competition between the two conflicting features: the smoothness and the fidelity. The earliest motivation to this approach dates back to 1923 when Whittaker (1923) introduced *graduation*, which is also one of the earliest works of nonparametric regression in the literature. By using the residual sum of squares $\sum_{t=1}^{N}(y_t - T_t)^2$ as the global measure of fidelity of the estimated trend T_t, Whittaker (1923) suggested to define the sum of the squares of kth order differences as the measure of roughness. Then the optimal trend is given by solving the following minimization scheme:

$$\text{(WH)}: \min_{\{T_t\}_{t=1}^{N}} \left\{ \sum_{t=1}^{N} (y_t - T_t)^2 + \lambda^2 \sum_{t=1}^{N-k} (\Delta^k T_t)^2 \right\}, \tag{3}$$

where order k and smoothing parameter λ are user-specified constants and Δ is the difference operator.

A particular example following Whittaker's approach is the HP filter (Hodrick and Prescott, 1997), which has been most extensively used in the real business cycle literature for detrending

$$\text{(HP)}: \min_{\{T_t\}_{t=1}^{N}} \left\{ \sum_{t=1}^{N} (y_t - T_t)^2 + \lambda \sum_{t=1}^{N} (T_{t-1} - 2T_t + T_{t+1})^2 \right\}. \tag{4}$$

Hodrick and Prescott recommended setting $\lambda = 1600$ when applying to real business studies. By manipulating the relevant first-order condition, the HP scheme leads to a two-way moving average with weights subjected to a damped harmonic approximately, see King and Rebelo (1993).

Recently, a popularly used measurement of the roughness penalty of estimates $T(x_t)$ in nonparametric regression is $\int [T''(x)]^2\,dx$. This leads to following penalized least-squares regression scheme:

$$\text{(SS)}: \min_{\{T_t\}_{t=1}^N} \left\{ \sum_{t=1}^{N} (y_t - T_t)^2 + \lambda \int [T''(x)]^2\,dx \right\}, \tag{5}$$

which is known as smoothing spline estimator. Remarkably, the problem of optimization SS over the space of all twice differentiable functions on the interval $[a,b]=[x_1,x_N]$ has a unique solution $T_\lambda(x)$ which is defined as the *natural cubic spline*, see Eubank (1999) and references therein.

More recently, Mosheiov and Raveh (1997) (MR) proposed a linear programming approach to estimate the trend by employing the sum of the *absolute* values rather than the common sum of squares to measure the smoothness and fidelity. The tradeoff balancing leads to such an optimization scheme

$$\text{(MR)}: \min_{\{T_t\}_{t=1}^N} \left\{ \lambda \sum_{t=1}^{N} |y_t - T_t| + (1-\lambda) \sum_{t=1}^{N-2} |T_{t+2} - 2T_{t+1} + T_t| \right\}. \tag{6}$$

Through a trick of variable changing, the objective function of the minimization scheme MR will be free of the absolute operator. However, extra constraints, monotonicity or polytonicity, have to be forced upon estimates in order to uniquely solve the linear programming problem. The location of the changing points of polytone trend is case-dependent and its selection is somewhat arbitrary.

Obviously, all these optimization schemes are closely related. The pointwise roughness measure of MR, $T_{t+2}-2T_{t+1}+T_t$, is essentially the same as those of HP and WH, and is the discrete version of SS. Instead of forward difference, $T_{t+2}-2T_{t+1}+T_t$, the backward and central difference approximation may also be used in the optimization MR. Moreover, with appropriate boundary modifications (such as in HP filter, see Baxter and King, 1995), the summation of the global smoothness measure in WH and MR can be processed from 1 to N, which may be more reasonable in a comparison with the integration in SS. Apart from their common motivation, another common feature of this class of nonparametric approaches is their use of global implicit minimization. As is well known, a global minimization problem can be quite expensive for its numerical computation.

The main objective of this paper is to present a new nonparametric approach, jump process, to estimate the unknown deterministic trend function of a time series. In contrast to the global implicit minimization approach, a localized approach is developed for trend estimation. The proposed jump process can locally minimize both characteristics and explicitly balance the fundamental tradeoff between them. A weighted average form of the jump process estimate is derived in the present paper, so that the implementation of jump process becomes extremely simple. The connection between the present jump process approach and the traditional trend estimation methods, as well as the DSP is discussed. It is found that the weight function of the jump process filter approaches the normal kernel, as the smoothing parameter increases.

This paper is organized as follows. The formalism of the jump process is introduced in Section 2. Numerical techniques regarding different implementations and boundary modifications are discussed in detail. The method is illustrated through applications to several real price series and one simulated example in Section 3. Conclusions are given in Section 4.

2. Theory and algorithm

2.1. Local measurement of smoothness and fidelity

A common feature of optimization schemes WH, HP, SS and MR is to minimize the linear combination of the *global measure* of fidelity and smoothness, while using a "smoothing parameter" to balance the tradeoff between fidelity and smoothness. There are three key aspects in the design of these nonparametric trend estimation approaches:

(1) Define measures for fidelity and smoothness,
(2) balance the tradeoff by employing a smoothing parameter,
(3) minimize the linear combination of two measures to achieve an estimate which is optimal in the sense of the given measures.

In the present study, an iterative jump process will be considered, which actually admits a simpler optimization approach for trend estimation

$$\begin{aligned} T_t^{M+1} &= T_t^M + R(T_{t-1}^M - 2T_t^M + T_{t+1}^M) \\ &= T_t^M + R\Delta^2 T_t^M \\ T_t^0 &= y_t, \quad t = 1,\ldots,N, \end{aligned} \tag{7}$$

where ratio R ($R > 0$) and iteration parameter M are user-specified constants. The second term on the right-hand side of iterative process (7), $T_{t-1}^M - 2T_t^M + T_{t+1}^M$, is the pointwise measure of smoothness. To have a better understanding of this iterative jump process, T_t^M of Eq. (7) is rewritten in terms of y_t:

$$T_t^M = y_t + R\sum_{k=0}^{M-1} \Delta^2 T_t^k, \quad t = 1,\ldots,N. \tag{8}$$

Further denote $v_t^{M-1} = \sum_{k=0}^{M-1} \Delta^2 T_t^k$, this yields

$$(y_t - T_t^M) + Rv_t^{M-1} = 0, \quad t = 1,\ldots,N. \tag{9}$$

It is obvious that the first term on the left-hand side of (9), $(y_t - T_t^M)$, is the local measure of the fidelity, while the second term, v_t^{M-1}, is the accumulative local measure of smoothness. At each step of the iteration, this process guarantees that the sum of the linear combination of the local deviation from y_t and the accumulative local measure of smoothness equals to zero. In the sense of such local minimization, the optimal estimated trend is the output of iterative jump process (7) by using y_t as input. Such a trend estimation method will be referred as a jump process estimator.

In terms of minimization, the relationship between the jump process and the optimization schemes WH, HP, SS, and MR is somewhat analogous to the relationship between the collocation and Galerkin approximation schemes well known in numerical analysis (see for example, Wei, 2000). The previous optimization schemes minimize the criterion function over the entire domain to obtain optimal estimates, while jump process forces the criterion function to pass through zero at each step of the iteration to give an optimal trend.

Besides the minimization of two properties, another important aspect of the construction of nonparametric trend estimation approach is the tradeoff balance. To illustrate how jump process (7) balances the tradeoff between smoothness and fidelity, the stability of jump process (7) is considered first. For this purpose, Eq. (7) is rewritten in a matrix form,

$$\mathbf{T}^{M+1} = \mathbf{A}\mathbf{T}^{M}, \tag{10}$$

where $\mathbf{T}^M = (T_1^M, T_2^M, \ldots, T_N^M)'$, and the tridiagonal matrix $\mathbf{A}$ has coefficients: $a_{t,t-1} = a_{t,t+1} = R$ and $a_{t,t} = 1 - 2R$, for $t = 1, 2, \ldots, N$. If all of the eigenvalues of $\mathbf{A}$ are smaller than unity, the iterative correction $\varepsilon^{M+1} = \|\mathbf{T}^{M+1} - \mathbf{T}^M\|$ will decay, then the process is stable. Since each diagonal term of the matrix is a constant, the eigenvectors of $\mathbf{A}$ can be represented in terms of a complex exponential form,

$$\mathbf{T}_t^M = q^M \mathrm{e}^{\mathrm{i}\gamma t}, \tag{11}$$

where $\mathrm{i} = \sqrt{-1}$ and γ is a wavenumber that can be chosen arbitrarily. Substituting Eq. (11) into Eq. (10) and removing the common term $\mathrm{e}^{\mathrm{i}\gamma t}$, an explicit expression for the eigenvalue q is obtained

$$q = 1 + 2R(\cos\gamma - 1). \tag{12}$$

For a stable process, the magnitude of this quantity is required to be smaller than unity,

$$q^2 = [1 + 2R(\cos\gamma - 1)]^2 < 1 \tag{13}$$

and q is maximum when $\cos\gamma = -1$. Therefore, the iterative process is stable provided $R < \frac{1}{2}$.

Thus, when $0 < R < \frac{1}{2}$, one has $\varepsilon^{M+1} \leqslant \varepsilon^M$, for any $M \in \mathbb{Z}^+$. Due to $\varepsilon^{M+1} = \|\mathbf{T}^{M+1} - \mathbf{T}^M\| = \|\Delta^2 \mathbf{T}^M\|$, is actually the global smoothness measure of estimated trend at the Mth iteration step, one can argue that as the iterative process is carried out longer and longer, the estimated trend becomes smoother and smoother, while the deviation of $\mathbf{T}^M$ from $\mathbf{Y} = (y_1, y_2, \ldots, y_N)'$ becomes larger and larger. Two smoothing parameters, R and M, govern the fundamental tradeoff between the smoothness and fidelity. In practice, R can be pre-fixed in the iteration process and only M is used to achieve the desired tradeoff.

2.2. Weighted average form of jump process

The advantage of the proposed approach is its simplicity, robustness and efficiency. However, the relationship between final estimates and original time series needs to be clarified for jump process (7). Fortunately, like most of the other nonparametric approaches, the estimated trend of the jump process also permits a weighted average representation in terms of the original series y_t. If M equals to 1, the estimates are

$$T_t^1 = Ry_{t-1} + (1-2R)y_t + Ry_{t+1}, \quad t = 1,\dots,N, \tag{14}$$

which is clearly a local weighted average form for y_t. In general, after M iterations, the estimated trend can be represented as

$$T_t^M = \sum_{k=t-M}^{t+M} W(k,M)y_k, \tag{15}$$

where weight function $W(k,M)$ has the general form of

$$W(k,M) = \begin{cases} \displaystyle\sum_{h=0}^{(M-k)/2} g(k,M,2h) & \text{when } M-k \text{ even,} \\ \displaystyle\sum_{h=1}^{(M-k+1)/2} g(k,M,2h-1) & \text{when } M-k \text{ odd,} \end{cases} \tag{16}$$

and

$$g(k,M,h) = \frac{M!R^{M-h}(1-2R)^h}{((M+k-h)/2)!((M-k-h)/2)!h!}. \tag{17}$$

It can be easily verified that,

$$\sum_{k=t-M}^{t+M} W(k,M) = 1 \tag{18}$$

and

$$W(-k,M) = W(k,M) \quad \forall k = 1,\dots,M. \tag{19}$$

Eq. (15) indicates that the jump process estimator can be viewed as a kernel smoother with (16) as a jump process kernel. From the point of view of the DSP, the weight function $W(k,M)$ is a low-pass filter. The implementation of the jump process becomes extremely simple due to the existence of (15). Therefore, the weighted average form (15) is very useful numerically.

The weight assignment of the jump process filter is analogous to that of other kernel regression methods. When M is not too small, and for any reasonable choice of R, such as $0.1 \leqslant R < 0.5$, the greater of the weights is assigned to the points close y_i, the smaller weight will be assigned to the points far away from y_i, see Table 1 and Fig. 1. It can also be seen from the figure that, when M is large, although the involved neighborhood is large, the effective window size of significant nonzero filter coefficients is smaller than $2M+1$.

A simple moving average filter can be formed by convolving $(\frac{1}{2}, \frac{1}{2})$ with itself $2M$ times. When $M = 1$, such a filter is the so-called Hanning filter (see Goodall, 1990)

$$(W_{-1}, W_0, W_1) = (\tfrac{1}{4}, \tfrac{1}{2}, \tfrac{1}{4}). \tag{20}$$

By setting $R = \frac{1}{4}$ in Eq. (14), it is clear that the one step jump process filter has the same coefficients as those of the Hanning filter and the M step jump process filter

Table 1
Weights of the jump process filter ($M = 6$)

k	$W_{k,6}$		
	General	$R = 0.4$	$R = 0.1$
0	$924R^6 - 1512R^5 + 1050R^4 - 400R^3 + 90R^2 - 12R + 1$	0.181824	0.390804
1	$-792R^6 + 1260R^5 - 840R^4 + 300R^3 - 60R^2 + 6R$	0.154368	0.227808
2	$495R^6 - 720R^5 + 420R^4 - 120R^3 + 15R^2$	0.12672	0.065295
3	$-220R^6 + 270R^5 - 120R^4 + 20R^3$	0.07168	0.01048
4	$66R^6 - 60R^5 + 15R^4$	0.039936	0.000966
5	$-12R^6 + 6R^5$	0.012288	0.000048
6	R^6	0.004096	0.000001

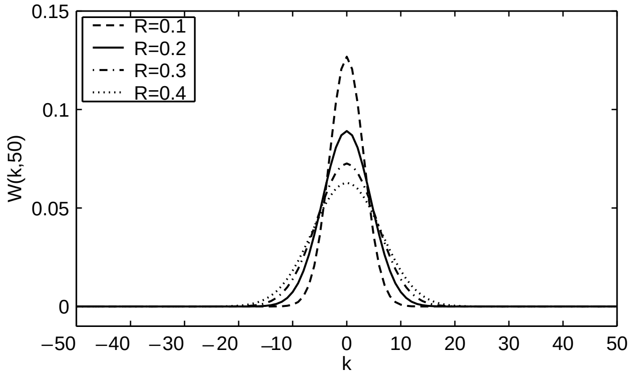

Fig. 1. Weights of the jump process filter ($M = 50$).

is identical to the digital filter obtained by convolving Hanning filter with itself M times. Thus, the proposed jump process filter can be viewed as a generalization of the Hanning filter.

2.3. Generalization of jump process estimator

For simplicity, only the uniformly spaced data have been considered for trend estimation so far. However, the framework of jump process estimator can be extended easily to randomly distributed data (i.e. nonuniformly spaced observations),

$$T_t^{M+1} = T_t^M + \bar{R}\,\frac{T_{t-1}^M(x_t - x_{t-1}) - T_t^M(x_{t+1} - x_{t-1}) + T_{t+1}^M(x_{t+1} - x_t)}{\frac{1}{2}(x_{t+1} - x_{t-1})(x_{t+1} - x_t)(x_t - x_{t-1})}$$

$$T_t^0 = y_t, \quad t = 1, \ldots, N. \tag{21}$$

This iteration process is stable provided $\bar{R} < (\Delta x)^2/2$, where $\Delta x = \min_t (x_t - x_{t-1})$. The weighted average form of (21) depends on the design. It will be quite complicated to obtain the explicit expression of the weighted function similar to (16). Fortunately, the regression estimates based on the iterative jump process (21) can be easily obtained. Therefore, the trend estimation based on jump process (21) will be useful even if either some data points are missing or a cross validation method is employed.

It is well known that the continuous counterpart of the jump process in stochastic processes is the diffusion process, which is usually represented in the form of a partial differential equation,

$$\frac{\partial T(x,\tau)}{\partial \tau} = \nabla^2 T(x,\tau),$$
$$T(x,0) = y(x). \tag{22}$$

Here, the temporal variable τ is the continuous time, rather than the time variable of the time series. To numerically simulate the diffusion process on uniformly spaced data, the second-order central difference and explicit Euler scheme may be employed for spatial and temporal discretizations

$$T_t^{M+1} = T_t^M + \frac{\Delta\tau}{(\Delta x)^2}(T_{t-1}^M - 2T_t^M + T_{t+1}^M),$$
$$T_t^0 = y_t, \quad t = 1,\ldots,N. \tag{23}$$

It is interesting to note that if one sets $R=\Delta\tau/(\Delta x)^2$, discretized approximation scheme (23) is the same as the iterative jump process (7) and the stability of explicit Euler scheme also requires mesh ratio $\Delta\tau/(\Delta x^2) < 1/2$. Thus, the diffusion Eq. (22) is capable of providing alternative perspective for the understanding of the jump process estimator. For example, the nonuniform jump process (21) can be easily derived from the heat Eq. (22). Obviously, high-order spatiotemporal discretization of Eq. (22) can also be used to construct a family of jump processes.

2.4. Jump process and Wiener process

For an appropriate range of R ($0 < R < 1/2$), the coefficients of (14), i.e. $(R, 1-2R, R)$, are nonnegative and $R + (1-2R) + R = 1$. Therefore, these coefficients may be interpreted as probabilities. Consider a particle at position k on the x-axis at time $\tau = M\Delta\tau$, in the next $\Delta\tau$ time period, the particle can have only three possible states: forward Δk, backward Δk, no change in position, with probabilities of P^+, P^- and P, respectively,

$$\Delta k = \begin{cases} \Delta x & \text{with probability } P^+ = R, \\ 0 & \text{with probability } P = 1-2R, \\ -\Delta x & \text{with probability } P^- = R, \end{cases} \tag{24}$$

where $k + \Delta k$ is the particle position after $\Delta\tau$. The process (24) is usually called a jump process in the stochastic process analysis, see for example Cox and Ross (1976). If a particle follows the jump process (24) and starts at the origin of the x-axis at $\tau = 0$, after M steps, it is easy to prove that the probability of this particle at position k is exactly the weight $W(k,M)$ given by (16). Hence, further investigation of jump process (24) will provide considerable insight into the proposed jump process filter.

It is obvious that the local mean and variance of Δk in (24) are

$$
\begin{aligned}
E\{\Delta k\} &= \Delta x(P^+ - P^-) = 0 \\
\mathrm{Var}\{\Delta k\} &= \Delta x^2(P^+ + P^-) - (E\{\Delta k\})^2 \\
&= 2\Delta x^2 R \\
&= 2\Delta\tau.
\end{aligned} \tag{25}
$$

In the continuum limit of an infinitesimally small step size, the discrete model (24) yields

$$
\mathrm{d}k = \sqrt{2}\,\mathrm{d}Z, \tag{26}
$$

where $\mathrm{d}Z$ is a standard Wiener process with $E\{\mathrm{d}Z\} = 0$, $\mathrm{Var}\{\mathrm{d}Z^2\} = \mathrm{d}\tau$. This implies that the $\mathrm{d}k$ is also a one-dimensional Wiener process (Brownian motion without the drift). Hence, the increase of particle movement during a relatively long period of time τ is given by

$$
k(\tau) - k(0) = \sum_{t=1}^{M} \varepsilon_t \sqrt{2\,\mathrm{d}\tau}, \tag{27}
$$

where the ε_t $(t = 1, 2, \ldots, M)$ are random numbers drawn from a standardized normal distribution. Consequently, it can be shown that $k(\tau) - k(0)$ is normally distributed with (Hull, 1999, Section 10.2)

$$
\begin{aligned}
&E\{k(\tau) - k(0)\} = 0, \\
&\mathrm{Var}\{k(\tau) - k(0)\} = \sqrt{2\tau}.
\end{aligned} \tag{28}
$$

Here $k(0) = 0$ and $k(\tau) = k$. This means that under the jump process (24), the particle movement will follow the normal distribution in the continuum limit $\Delta\tau \to 0$. Since $\Delta\tau \to 0$ is equivalent to $M \to \infty$ when τ is fixed, and the particle movement probability function is exactly the weight function $W(k, M)$. One can conclude that the weight function $W(k, M)$ of the proposed jump process filter will approach the normal kernel at the limit of $M \to \infty$.

It is well known that the filter coefficients generated by convolving Hanning moving average filter approximate the Gaussian kernel as $M \to \infty$. The above finding indicates that, the present generalized Hanning filter, the jump process filter, shares the same property. Such property endows the jump process weight function to be a good kernel function for kernel regression, for which a widely used kernel is the Gaussian density. The proposed jump process weight function provides a *discrete approximation* to the Gaussian kernel. On the other hand, as pointed out in the DSP literature (such as Koenderink, 1984; Hummel, 1987), the solution of the heat diffusion Eq. (22) may equivalently be viewed as the result obtained by convolving original signal with the Gaussian kernel. This again agrees with the present finding about the jump process estimator.

2.5. Implementation particulars

It follows from above discussion that there are two simple and controllable ways to implement the jump process trend estimation. One way is based on iterative jump process (7). The time series $\{y_t\}$ is used as the input data. The iteration number M is used to control the final estimates. The other way is to use the weighted average form (15). The trend is estimated by convolving y_t and the jump process kernel once. In both ways, the R can be fixed and only M needs to be adjusted. Theoretically, the estimated trends from two ways are the same, however, there are some minor differences due to possible different boundary treatment and applicability. Generally speaking, for uniform spacing data trend estimation problems, the convolution implementation is more efficient than iterative implementation. However, the iterative implementation can be easily done for randomly spaced data regression, for which the weight kernel of the form of (15) is difficult to be constructed.

Another difference of these two implementations is the different possible modification in dealing with boundary effect. The boundary effect is a common thorny problem for linear filtering and kernel smoothing, i.e., linear symmetric filters fail to provide estimates for the initial or/and end terms of the series (Kendall et al., 1983). The problem seems to be more serious in a convolution, since a larger computational support will locate outside the boundaries in this case, while there is only one point outside each boundary for the iterative way at each iteration step.

In the literature, there are some alternative approaches for dealing with boundary effect (Goodall, 1990). One approach uses progressively more asymmetric versions of filters at the end points, it will result in more biased estimates. Such techniques were widely used in moving average filtering and kernel smoothing, see for example Gasser and Müller (1979), and can be directly adopted by the convolution implementation of the jump process estimation. The counterpart of such a technique in an iterative implementation is the well-known upwind difference approximation scheme in numerical analysis. However, though based on the same motivation, these two modifications along with two implementations generally yield different estimates.

Artificially "padding data" or extrapolating the series is another approach for generating necessary support for the symmetric filter. According to the observed characteristic of trend component, repeating the latest observation, symmetric or antisymmetric extension may be used. The same extension technique, such as symmetric or antisymmetric, can be adopted by both convolution and iterative implementations. Furthermore, it can be proved that by using the same symmetric/antisymmetric extension technique, the final estimates of two different implementations of the jump process are identical. Therefore, we limit our attention in the present study to the implementation of convolution with boundary extensions.

3. Numerical experiments

3.1. Real time series analysis

The effect of varying the smoothing parameter M on the estimated trend is studied. A real time series, the 'Sales of Company X' series, is employed in the present work. Such a series is a well-known test case, see Chatfield and Prothero (1973), Box and

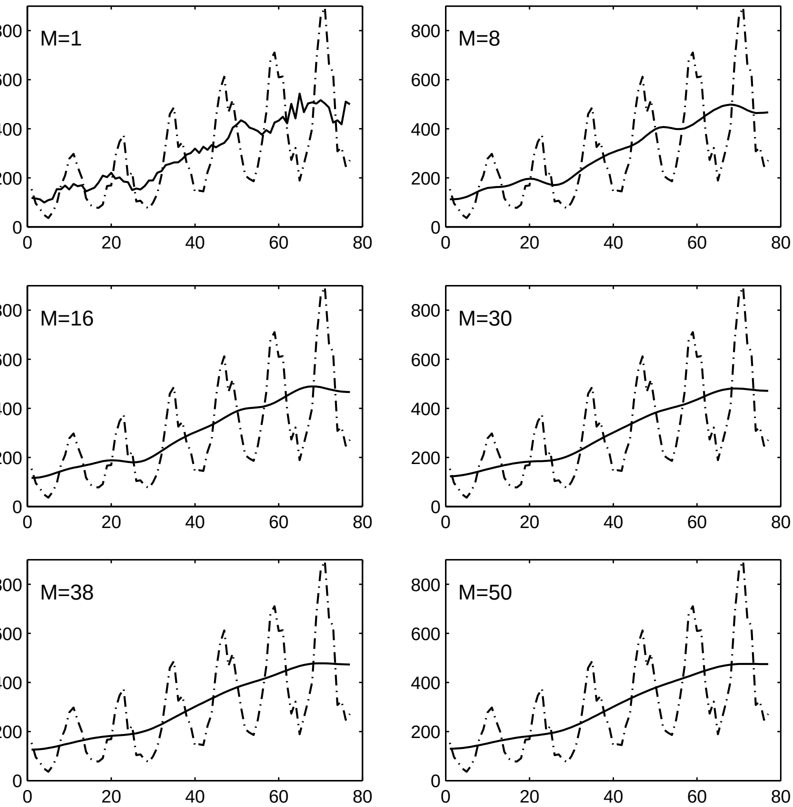

Fig. 2. Trend estimated by using the jump process as a function of M. The Chatfield–Prothero case study: "Sales of Company X".

Jenkins (1973), and Mosheiov and Raveh (1997). This test case is a monthly series ranging from January 1965 to May 1971. The time series is relatively regular, with a monotonic growing trend and clearly identifiable seasonal component. Seasonally adjusted series, which is obtained by using the X-12-ARIMA of Findley et al. (1998), is used to convolve with the jump process weight kernel, and the symmetric extension is employed at boundaries.

The value of M is varied from 1 to 50 with a fixed value of R ($R = 0.4$). As one may expect, estimates with large smoothing parameters are very smooth, while for small values of M the estimates still contain high frequency components, see Fig. 2. It can be seen clearly from Fig. 2, the slope of the trend changes around the 28th observation. This agrees with the finding by Mosheiov and Raveh (1997). However, when M becomes larger and larger, the estimated trend fails to provide relatively reasonable tendency at the right boundary due to the boundary effect. Two other boundary modifications are also tested in this case study. One is to multiply the jump process kernel by appropriate linear functions to generate progressively more asymmetric kernels near the boundaries, see Gasser and Müller (1979) for the details about such a modification. Another is to use the upwind difference approximation along with an iterative implementation. However, these results are not as good as the present symmetric extension. We argue that one cannot look forward to obtaining better results by using other boundary modifications, since, all modifications rely on a one-sided approximation, which inherently introduces serious bias.

For a pre-fixed R, the choice of fundamental tradeoff indicator M is obviously a crucial decision. The optimal M is clearly determined by the signal to noise ratio in the data. However, the practical choice of R is usually dominated by the preferences of the practitioner or by the question being asked. The visually best principle is used as the guideline to estimate the trend here and such an estimated trend reflects our natural preference for the trend, viz. the estimates are smooth enough to indicate some long-run tendency. In this test case, $M = 30$ is chosen as the approximate optimal smoothing parameter.

To validate the proposed jump process estimator further, a well studied economic series, the Spanish Industrial Production Index for Energy (see Ferreira et al., 2000), is employed. The time series consists of monthly data of the IPI for energy from January 1975 up to December 1993. Although this series is somewhat more irregular than the previous studied series, nonmonotonic trend and seasonal components are still identifiable from time series, see Fig. 3(a). Again, the seasonally adjusted series by the X-12-ARIMA is used as input.

The X-12-ARIMA (Findley et al., 1998) is the most commonly used method of seasonal adjustment for official statistics throughout the world. Besides seasonal adjustment, it also can present trend estimates by employing the Henderson filter (see Kenny and Durbin, 1982; Blanchi et al., 1999). For the IPI series, the trend estimated by the X-12-ARIMA with default *variable trend cycle routine* is shown in Fig. 3(a). Visually, this estimate is not smooth enough to indicate some long-run tendency. This suggests that the band-width of the Henderson filter should be increased. To study estimates of large band-width, the symmetric Henderson filter is independently implemented in the present study. For a comparison, the implementation of the Henderson filter is the same as that of the jump process filter, i.e, the antisymmetric and symmetric extensions are employed at left and right boundaries, respectively. The trend estimates of both the Henderson and jump process filters are also depicted in Fig. 3(a). It is clear that the results of two estimates are almost the same, and capture the underlying tendency.

The estimated trend, $\hat{T}_t$, allows us to estimate the annual underlying growth of the Spanish IPI for energy, which is an issue of interest in economics. The estimate to underlying growth, c_t, can be calculated by accumulating the last 12 basic growths (Ferreira et al., 2000),

$$\hat{c}_t = [L^0 + L^1 + \cdots + L^{11}]\Delta\hat{T}_t, \tag{29}$$

where L denotes a lag operator. The estimates of the underlying growth based on the previous estimated trends are plotted in Fig. 3(b). It can be seen from the figure that the estimate of the X-12-ARIMA is not a smooth curve, and oscillates quite frequently, while the estimates of the Henderson and the jump process filters can indicate a pattern

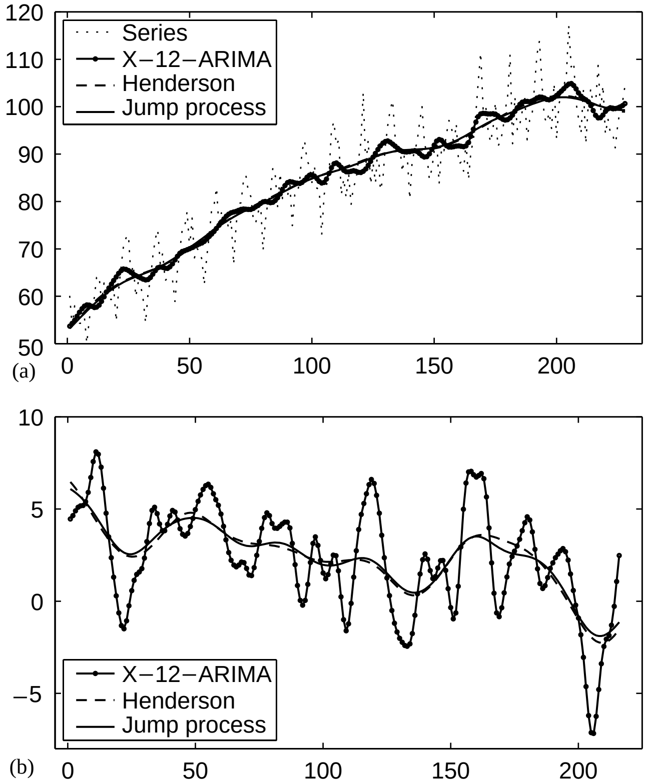

Fig. 3. Estimation of the Spanish IPI for energy: (a) estimated trend by using the X-12-ARIMA, Henderson filter ($M = 40$), and jump process filter ($R = 0.4$ and $M = 100$), (b) estimated annual underlying growth.

which is similar to that of Ferreira et al. (2000), thus could be interpretable from the point of view of economics.

Finally, the jump process filter is tested by using another well studied case. This time series consists of the Beveridge index (Beveridge, 1921) of wheat prices from the year 1500–1869 (Anderson, 1971; Hart, 1991, 1994). These data are an annual index of prices at which wheat was sold in European markets. To correct for heteroscedasticity in the original series, a logarithmic transformation is employed. The chart (Fig. 4) clearly indicates that the transformed series consists of a nonstationary trend and a constant variance random noise. There is no regular cyclical component presented in time series. The estimated results are depicted in Fig. 4. Here the smoothing parameters are chosen as $M = 48$ for the Henderson filter, and as $R = 0.4$ and $M = 120$ for the jump process filter. The symmetric extension is used at boundaries for both filtering. It can be seen from Fig. 4, estimates of two filters are almost identical, and describe the most slowly changing part of the series.

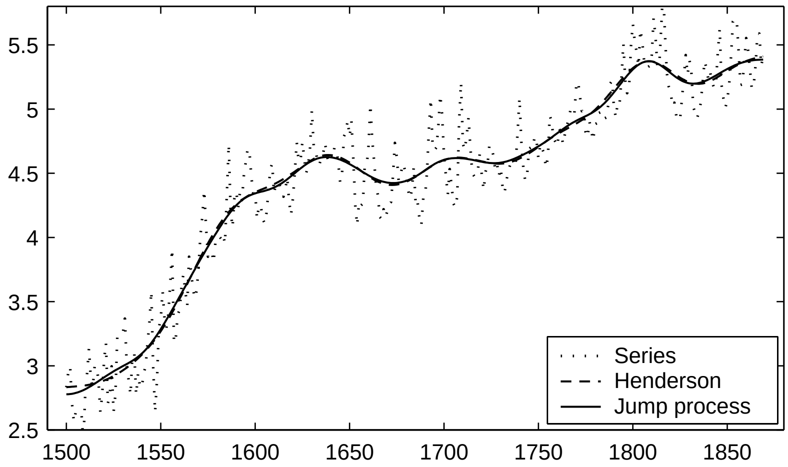

Fig. 4. Trend estimated by using the Henderson filter ($M = 48$) and the jump process filter ($R = 0.4$ and $M = 120$) for the Beveridge wheat price index data.

3.2. Simulation study

For a quantitative test of the proposed jump process estimator, a simulation study is considered. To this end, a simple model is constructed which is analogous to the additive model in the absence of seasonal component

$$y(x) = T(x) + \varepsilon(x), \tag{30}$$

where $T(x)$ is a polynomial trend given by (Hart and Wehrly, 1986; Høst, 1999)

$$T(x) = \mu + 10x^3 - 15x^4 + 6x^5, \tag{31}$$

which mimics the typical behavior of growth curves. Here estimates will be calculated on a uniform mesh $x_t = t/N$, for $t = 1, \ldots, N$, in the interval of $[0, 1]$. The random white noise $\varepsilon(x_t)$ is normally distributed with zero mean and a constant variance σ^2. The value of μ is set to 1 to ensure that $y(x_t)$ is positive. Both the simulated data and the underlying trend are shown in Fig. 5.

Investigation is numerically carried out first on how the smoothing parameter M governs the fundamental tradeoff when R is pre-fixed in the jump process estimation. One series is simulated with 101 uniform locations ($N = 101$) and $\sigma = 0.1$ for this purpose. The symmetric extension is used throughout the simulation study. By varying M from 1 to 100, different estimates are obtained. The following global measures are utilized to indicate the corresponding characteristics of estimated trend $\hat{T}_t$

$$\text{Smoothness } \{\hat{T}_t\}\text{: } \sqrt{\frac{1}{N}\sum_{t=1}^{N}(\Delta^2 \hat{T}_t)^2}, \tag{32}$$

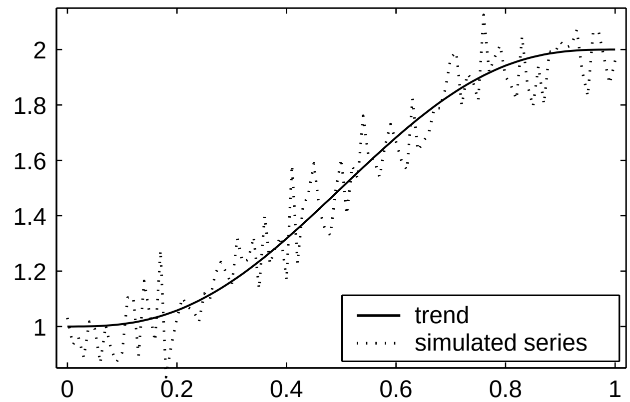

Fig. 5. Plot of simulated series and trend component (noise standard deviation $\sigma = 0.1$ and sample size $N = 101$).

$$\text{Fidelity } \{\hat{T}_t\}: \sqrt{\frac{1}{N}\sum_{t=1}^{N}(y_t - \hat{T}_t)^2}, \tag{33}$$

$$\text{Bias } \{\hat{T}_t\}: E\{|T_t - \hat{T}_t|\}. \tag{34}$$

The results are shown in Fig. 6. These numerical results verify our aforementioned theoretical investigation, viz. when R is pre-fixed, M governs the smoothness-fidelity tradeoff and the bias-variance tradeoff.

The problem of optimal trend estimation is considered next. With a constructed trend, mean square error (MSE) is employed to measure the deviation of the numerically estimated trend from the true trend

$$\text{MSE} = \frac{1}{N}\sum_{t=1}^{N}(T_t - \hat{T}_t)^2. \tag{35}$$

Three noise levels, $\sigma = 0.025, 0.05, 0.1$, and three sample sizes, $N = 51, 101$, and 201 are investigated. For each combination of N and σ, 100 independent sets of data are generated. Both the jump process and Henderson filters are tested. Goodness-of-fit tests are employed to search for an optimal smoothing parameter which minimizes MSE in all cases. The results of simulation study are summarized in Table 2. As can be seen from the table, the average value of the minimal MSE is smaller than 1.0×10^{-3} in all cases. In other words, the satisfactory trend estimates can be obtained by using both the Henderson and jump process filters in all parameter combinations. The optimal estimated trend is actually very close to the true underlying trend, see example plots in Fig. 7. Another observation from Table 2 is that the optimal MSE of the Henderson estimates is commonly smaller than those of the jump process estimates. However, such difference is very small, especially when sample size N is large. Visually, the difference between the estimates of the Henderson and jump process filters is negligible when $N = 201$, see Fig. 7.

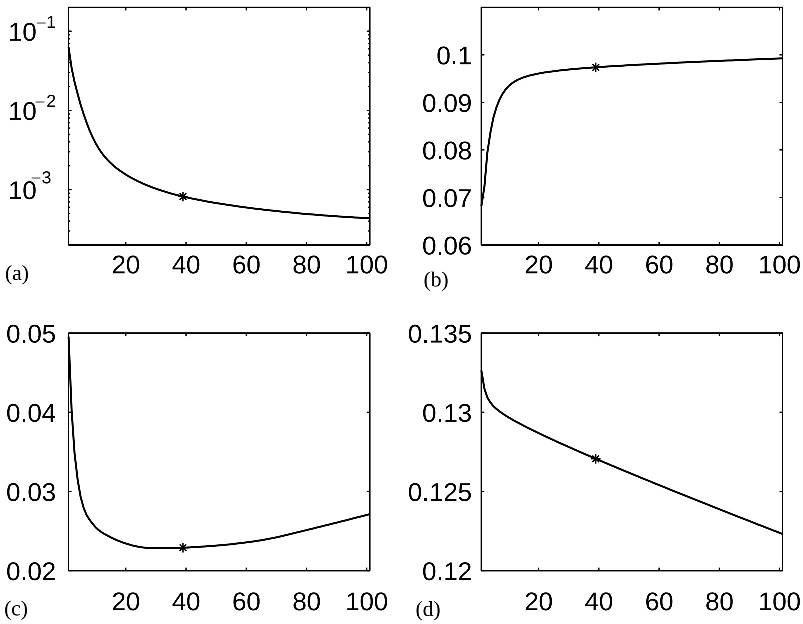

Fig. 6. The plots of the characteristics of estimated trend vs the smoothing parameter M ($\sigma = 0.1$, $N = 101$ and $R = 0.3$): (a) smoothness; (b) fidelity; (c) bias; (d) variance. The minimizer of MSE is $M = 39$ for this set of data, and the relevant location is marked by $*$ in the plots. The log-scale plot is used for (a) since smoothness measure decays very quickly.

Table 2
The results of simulation study. Here MSE ($\times 10^5$) is the average value of the minimal MSE. The corresponding average value of the optimal smoothing parameter M is in parentheses. In the jump process estimator, $R = 0.45$

Filter	σ	MSE		
		$N = 51$	$N = 101$	$N = 201$
Jump process	0.025	10.18(7.02)	5.58(19.39)	3.06(57.51)
	0.05	29.83(11.62)	16.87(34.99)	9.27(101.28)
	0.1	89.83(20.45)	50.56(63.02)	28.25(177.39)
Henderson	0.025	6.22(17.18)	3.26(33.84)	1.79(61.57)
	0.05	20.76(22.39)	10.77(41.78)	6.03(75.27)
	0.1	63.80(29.79)	34.59(53.36)	20.11(93.50)

It is clear from Table 2 that the optimal MSE and smoothing parameter M change when N or σ is different. Further investigation is needed to clarify the pattern of these changes. The smoothing parameter M provides information about the signal-to-noise ratio (SNR) in data. By keeping the trend component unchanged, when the noise devi-

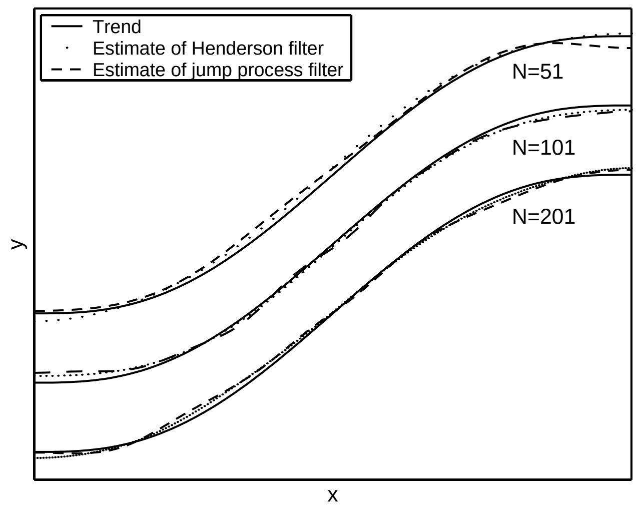

Fig. 7. Example plots of optimal estimated trends. In all plots, $\sigma = 0.1$, and $R = 0.45$ in the jump process filter.

ation σ increases from 0.025 to 0.1, the SNR also increases. Correspondingly, a larger M is required to suppress the increasing noise, see Fig. 8(b). Since unwanted noise becomes larger, the relevant optimal MSE of trend estimates also increases, see Fig. 8(a). The increasing ratios displayed in Fig. 8 are clearly independent of the sample size N, since these ratios are the same for different N. Moreover, the increasing ratios of the jump process estimates are different from those of the Henderson estimates. In terms of the optimal MSE (Fig. 8(a)), the increasing ratios of the jump process estimates are higher than that of the Henderson estimates. Hence, although the MSE of the jump process estimates is slightly bigger than that of the Henderson estimates, the difference between them will become smaller and smaller as noise increases. In terms of the optimal smoothing parameter M (Fig. 8(b)), the increasing ratios of the jump process estimates are also higher than that of the Henderson estimates. This means that the optimal M of the jump process estimates is usually larger than that of the Henderson estimates. It is noted that the smoothing parameter M of the jump process can be relatively large, or even larger than the sample size N, since in the iterative jump process implementation, M is just an iteration number. Therefore, the need for a large M does not affect the usefulness of the jump process filter. However, by using a symmetric linear filter (such as Henderson filter) without invoking asymmetric versions, it could be somewhat cumbersome to deal with a finite sample trend estimation problem with very large noise, since the optimal band-width M might be quite large.

Another alternative way for analyzing the numerical results in Table 2 is to examine how the optimal MSE and relevant minimizer M change when the sample size N increases from 51 to 201. Such a numerical analysis is meaningful since it can describe

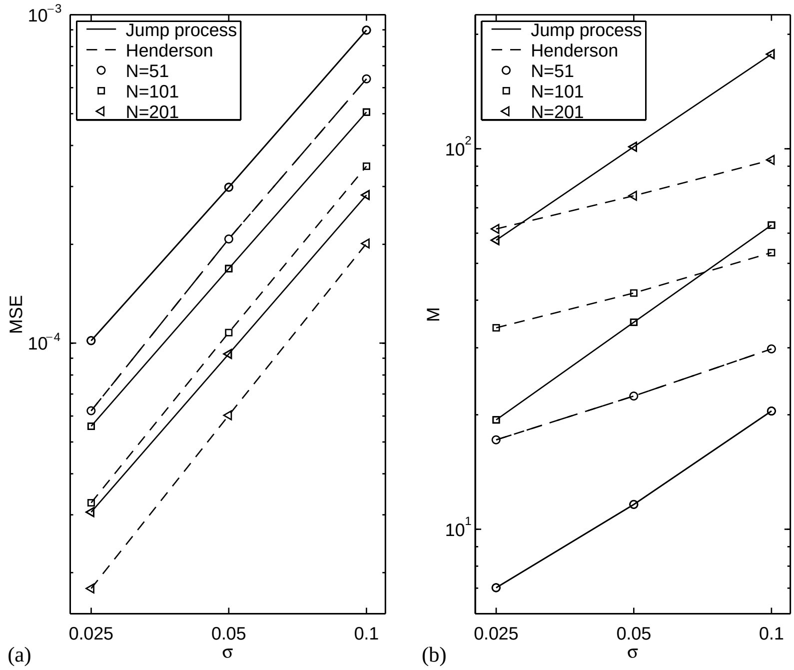

Fig. 8. (a) The log-scale plot of the average optimal MSE vs. noise standard deviation σ; (b) The log-scale plot of the average optimal M vs. σ.

the asymptotic behavior of the jump process estimates. It is known that the highest noise frequency in the discrete data is proportional to $\pi/\Delta x = \pi(N-1)$. On the other hand, the trend is a smooth function so that its frequencies do not change too much in the Fourier domain when N changes. Hence, when the sample size N increases, there are many more high frequencies in the Fourier domain which need to be eliminated. Consequently, a larger band-width is required to achieve optimal estimate, see Fig. 9(b). The numerical study reveals that the optimal estimates of both the jump process and Henderson filters converge quite rapidly to the true trend as sample size N becomes larger. Furthermore, the decreasing ratios of the jump process and Henderson estimations are almost identical, see Fig. 9(a). Again, the increasing tendency in Fig. 9(b) suggests that the optimal band-width M of the Henderson filter becomes very large when N is huge, and the optimal iteration number M of the jump process filter is usually even larger.

4. Conclusions

The main purpose of this paper is to introduce a jump process approach to the trend estimation of time series. The proposed jump process estimator explicitly minimizes

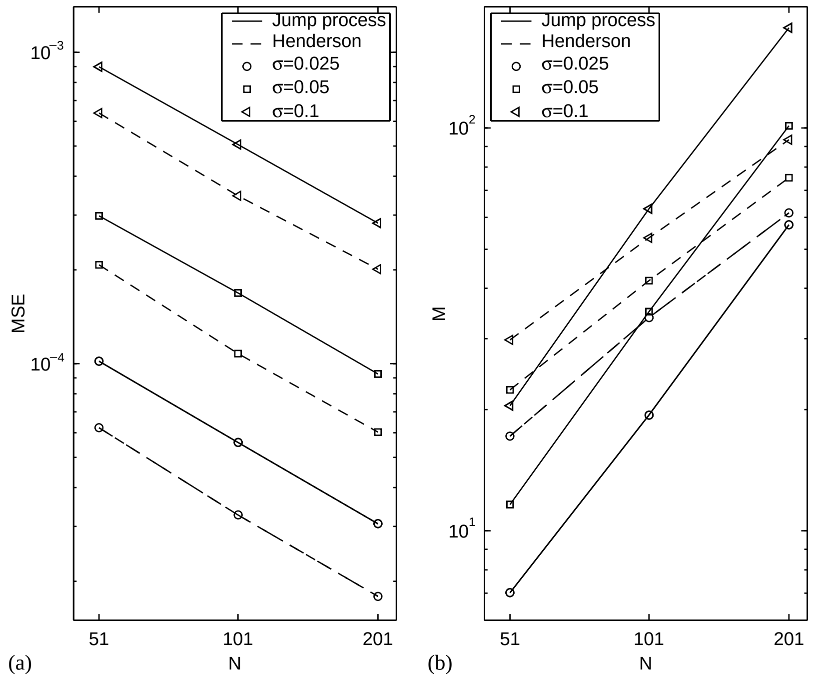

Fig. 9. (a) The log-scale plot of the average optimal MSE vs. sample size N, (b) The log-scale plot of the average optimal M vs. N.

the local smoothness and fidelity of a time series. A weighted average form of the jump process estimates is derived. The connection of the proposed approach to the Hanning filter, Gaussian kernel regression, the heat equation and the Wiener process is discussed in detail. The proposed method is validated by using both real data sets and a simulated time series.

Several trend estimation methods were successfully developed in the literature to quantify the competition between the smoothness and fidelity globally. In contrast, a jump process, which can locally minimize both features and explicitly balance the fundamental tradeoff between them, is proposed. The investigation opens up the opportunity for developing other new trend estimation methods which are optimal in a local sense. The feasibility, property, advantage and disadvantage of the trend estimation method development based on local optimization deserves further investigations.

The characteristic of the jump process estimator is its local optimization. In terms of minimization, the relationship between the jump process approach and traditional methods is analogous to the relationship between the collocation and Galerkin approximation in numerical analysis. The numerical strengths of the jump process are simplicity and robustness.

Like many nonparametric approaches, the estimate of the jump process also permits a weighted average form, which is very useful numerically. The weight shape of the jump process filter is also analogous to that of other kernel regression schemes. The present study reveals that the jump process filter can be viewed as a generalization of the Hanning filter. A jump process trend estimation scheme is developed for nonuniformly spaced data. Such a scheme is useful in case where some data points are missing or cross validation method is used. Furthermore, it is shown that the proposed jump process is equivalent to a discretized form of the heat diffusion equation. Hence the continuous diffusion process can be used as the other convenient starting point for theoretical analysis.

By examining the jump process from the point of view of stochastic process analysis, it is shown that such a jump process, in an appropriate limit, is a Wiener process. Hence, the weight function of jump process filter approaches the normal curve when smoothing parameter tends to infinity. This agrees with the relevant findings of the convolution Hanning filter, as well as the findings of the heat equation in the DSP literature.

Several real time series and a simulated test example are studied to demonstrate the usefulness of the jump process estimator. Although studied series varies from regular to irregular, the jump process performs very well in all cases. The numerical results of the jump process are in good agreement with both the present theoretical analysis and the results of previous studies.

The Henderson filter, which is embedded in the most popular seasonal adjustment package, X-12-ARIMA for trend estimation, is independently implemented to further validate the jump process filter. Even though the estimation error of the jump process is found to be a little bit larger than that of the Henderson filter for a particular case, both errors are actually small and the difference between them can be negligible for real applications. Moreover, numerical results reveal that such difference will decay when noise is larger. In fact, there is no much visual difference between the estimated trends of two methods in all example plots. On the other hand, numerical analysis from the perspective of asymptotic behavior indicates that the optimal estimate of the jump process is able to converge rapidly to true trend as the sample size becomes larger, with the same convergence speed as the Henderson filter. Comparatively speaking, the jump process filter is almost as good as the Henderson filter in terms of the performance of trend estimation for a given time series.

Numerical analysis also indicates that the optimal smoothing parameter M of two filters will increase when the sample size is large or the noise is large. Moreover, the optimal M of the jump process estimator is usually larger than that of the Henderson filter. In terms of iterative implementation of the jump process, M is the iteration number. At each iteration step, there is only one point outside boundaries, which need to be "predict" based on interior nodes. Consequently, it is applicable to using a quite large M in the jump process trend estimation. However, if the Henderson filter is kept in its symmetric form without invoking asymmetric versions, the boundary effect might be serious when a large band-width M is required. Since there are two simple and controllable ways to implement the jump process trend estimation, there are actually more freedom for the jump process estimator to deal with boundary effect. Therefore, the jump process trend estimation is numerically more flexible than the Henderson filter, especially in terms of dealing with the boundary effect.

Acknowledgements

This work was supported in part by the National University of Singapore.

References

Anderson, T.W., 1971. The Statistical Analysis of Time Series. Wiley, New York.

Ball, M., Wood, A., 1996. Trend growth in post-1850 British economic history: the Kalman filter and historical judgment. Statistician 45, 143–152.

Baxter, M., King, R., 1995. Measuring business cycles: approximate band-pass filters for economic time series. NBER Working Paper 5022.

Beveridge, W.H., 1921. Weather and harvest cycles. Econom. J. 31, 429–452.

Blanchi, M., Boyle, M., Hollingsworth, D., 1999. A comparison of methods for trend estimation. Appl. Econom. Lett. 6, 103–109.

Borra, S., Ciaccio, A.D., 2002. Improving nonparametric regression methods by bagging and boosting. Comput. Statist. Data Anal. 38, 407–420.

Box, G.E.P., Jenkins, G.M., 1973. Some comments on a paper by Chatfield and Prothero and on a review by Kendall. J. Roy. Statist. Soc. Ser. A 136, 337–352.

Campbell, J.Y., Perron, P., 1991. Pitfalls and opportunities: what macroeconomists should know about unit roots. NBER Macroeconom. Ann. 6, 141–201.

Canova, F., 1998. Detrending and business cycle facts. J. Monetary Econom. 41, 475–512.

Chatfield, C., 1996. The Analysis of Time Series: An Introduction. Chapman & Hall, London.

Chatfield, C., Prothero, D.L., 1973. Box-Jenkins seasonal forecasting: problems in a case study J. Roy. Statist. Soc. Ser. A 136, 295–315.

Cleveland, W.S., 1979. Robust locally weighted regression and smoothing scatterplots. J. Amer. Statist. Assoc. 74, 829–836.

Cochrane, J.H., 1991. Comment to pitfalls and opportunities: what macroeconomists should know about unit roots. NBER Macroeconom. Ann. 6, 201–210.

Cox, J., Ross, S., 1976. The valuation of option for alternative stochastic processes. J. Finance Econom. 3, 145–166.

Doksum, K., Koo, J.Y., 2000. On spline estimators and prediction intervals in nonparametric regression. Comput. Statist. Data Anal. 35, 67–82.

Donoho, D.L., Johnstone, I.M., Kerkyacharian, G., Picard, D., 1995. Wavelet shrinkage: asymptopia? J. Roy. Statist. Soc. Ser. B 57, 301–369.

Eubank, R.L., 1999. Nonparametric Regression and Spline Smoothing. Marcel Dekker, New York.

Fan, J., Gijbels, I., 1996. Local Polynomial Modelling and its Applications. Chapman & Hall, London.

Ferreira, E., Núñez-Antón, V., Rodríguez-Póo, J., 2000. Semiparametric approaches to signal extraction problems in economic time series. Comput. Statist. Data Anal. 33, 315–333.

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. Business Econom. Sat. 16, 127–177.

Franses, P.H., 1998. Time Series Models for Business and Economic Forecasting. Cambridge University Press, Cambridge.

Gasser, T., Müller, H.G., 1979. Kernel estimation of regression functions. In: Gasser, T., Rosenblatt, M. (Eds.), Smoothing Techniques for Curve Estimation. Springer, Berlin, pp. 23–68.

Goodall, C., 1990. A survey of smoothing techniques. In: Fox, J., Long, J.S. (Eds.), Modern Methods of Data Analysis. Sage Publications, Newbury Park, CA, pp. 126–176.

Härdle, W., Tuan, P.D., 1986. Some theory on M-smoothing of time series. J. Tim. Ser. Anal. 7, 191–204.

Hart, J.D., 1991. Kernel regression estimation with time series errors. J. Roy. Statist. Soc. Ser. B 53, 173–187.

Hart, J.D., 1994. Automated kernel smoothing of dependent data by using time series cross-validation. J. Roy. Stat. Soc. Ser. B 56, 529–542.

Hart, J.D., Wehrly, T.E., 1986. Kernel regression estimation using repeated measurement data. J. Amer. Statist. Assoc. 81, 1080–1088.

Harvey, A.C., 1989. Forecasting, Structural Time Series Models and the Kalman Filter. Cambridge University Press, Cambridge.

Hodrick, R.J., Prescott, E.C., 1997. Postwar US business cycles: an empirical investigation. (Carnegie Mellon University, 1980) J. Money Credit Banking 29(1), 1–16.

Høst, G., 1999. Kriging by local polynomials. Comput. Statist. Data Anal. 29, 295–312.

Hull, J.C., 1999. Options, Futures, and Other Derivatives. Prentice-Hall, Upper Saddle River, NJ.

Hummel, A., 1987. Representations based on zero-crossing in scale-space. Proceeding IEEE Computer Vision and Pattern Recognition Conference, pp. 204–209. In: Fischler, M.A., Firschein, O. (Eds.), Readings in Computer Vision: Issues, Problems, Principles and Paradigms. Morgan Kaufmann, Los Altos, CA.

Keilegom, I.V., Akritas, M.G., Veraverbeke, N., 2001. Estimation of the conditional distribution in regression with censored data: a comparative study Comput. Statist. Data Anal. 35, 487–500.

Kendall, M.G., Stuart, A., Ord, J.K., 1983. The Advanced Theory of Statistics, Vol. 3. Griffin, London.

Kenny, P.B., Durbin, J., 1982. Local trend estimation and seasonal adjustment of economic and social time series. J. Roy. Statist. Soc. Ser. A 145, 1–41.

King, R.G., Rebelo, S.T., 1993. Low frequency filtering and real business cycles. J. Econom. Dynamics Control 17, 207–231.

Koenderink, J.J., 1984. The structure of images. Biol. Cybernet. 50, 363–370.

Meade, N., Islam, T., 1995. Prediction intervals for growth curve forecasts. J. Forecasting 14, 413–430.

Mills, T.C., Crafts, N.F.R., 1996. Modelling trends in economic history. Statistician 45, 153–159.

Mosheiov, G., Raveh, A., 1997. On trend estimation of time series: a simple linear programming approach J. Oper. Res. Soc. 48, 90–96.

Müller, H.-G., 1988. Nonparametric Regression Analysis of Longitudinal Data. . Lecture Notes in Statistics, Vol. 46. Springer, Berlin.

Nelson, C.R., Plosser, C., 1982. Trends and random walks in macroeconomic time series. J. Monetary Econom. 10, 139–162.

Pollock, D.S.G., 2000. Trend estimation and de-trending via rational square-wave filters. J. Econometrics 99, 317–334.

Sims, C.A., 1988. Bayesian scepticism on unit root econometrics. J. Econom. Dynamics Control 12, 463–474.

Sims, C.A., Uhlig, H., 1991. Understanding unit rooters: a helicopter tour Econometrica 59, 1591–1599.

Visser, H., Molenaar, J., 1995. Trend estimation and regression analysis in climatological time series: an application of structural time series models and the Kalman filter J. Climate 8, 969–979.

Wei, G.W., 2000. A unified approach for the solution of the Fokker–Planck equation. J. Phys. A-Math. Gen. 33, 4935–4953.

Wen, Y., Zeng, B., 1999. A simple nonlinear filter for economic time series analysis. Econom. Lett. 64, 151–160.

Whittaker, E.T., 1923. On a new method of graduation. P. Edinburgh Math. Soc. 41, 63–75.